# PREFACE

## TO FIFTH EDITION

THE field of radio communication is continually widening and with it the form and scope of examinations on the subject. Moreover, the prospect of continued development makes it increasingly impracticable to attempt to cover the whole sphere in any one book.

The scope of the present work, therefore, together with Volume I, has been limited to the fundamental groundwork of radio communication, and it is hoped that it will not only continue to serve as a handbook for the practical radio engineer but will also cover the greater part of the latest City and Guilds syllabus in Communications Engineering, in respect of Telecommunications Principles, I to V, and Radio, I to IV.

The relatively small field which is not covered is that in which the rapid developments of to-day are occurring, namely in short-wave and microwave operation, and in the technique of the cathode-ray oscillograph. These fields are dealt with extensively in the companion volumes *Short Wave Radio* and *Cathode Ray Oscillographs* and in the present work brief mention only has been made of the latest developments, leaving the reader to refer for a more detailed treatment to one or other of these volumes.

Certain additional information has been included, particularly in connexion with transmission lines and network theory, and by courtesy of the City and Guilds of London Institute examination papers for 1945 and 1946 have been included at the end.

J. H. REYNER

BOREHAM WOOD
*March*, 1947

# PREFACE

THIS book is intended to provide a concise résumé of modern radio engineering practice for the more advanced student. The fundamental principles outlined in Volume I have been amplified, and extra material has been added to make the work reasonably complete. The general scope of the work is such as to cover the final stage of the City and Guilds Examination, the Preliminary and Intermediate Grades being dealt with by the first volume, to which this present book must be considered as an adjunct.

The non-mathematical character of the work has been maintained, but a particular effort has been made to provide a thorough groundwork of fundamental ideas. Hence, detailed proofs of some of the more important statements are included, but extensive treatment of particular cases is not provided, as the reader should be able to achieve this for himself.

Where possible, every statement has been explained. This applies particularly to the treatment of Feeders, Coupled Circuits, and Filter Networks, subjects which are usually treated either intensely mathematically or with a vagueness which still leaves the reader mystified as to what really takes place.

As a result it is hoped that, despite the necessary limitation of space, this book with its companion will provide a complete groundwork for the Radio Engineer.

J. H. REYNER

BOREHAM WOOD
*January*, 1936

# CONTENTS

## CHAPTER XI



## CHAPTER XII



## CHAPTER XIII



## CHAPTER XIV



## CHAPTER XV

## CHAPTER XVI



## CHAPTER XVII



## APPENDIX

# RADIO COMMUNICATION

## CHAPTER I

### THE RADIO TRANSMITTER

RADIO transmission to-day is carried out practically entirely by valve transmitters. Arcs and high-frequency alternators have been ousted and the spark transmitter only remains in low-power ship sets. Even here it is being displaced by tonic train and radio telephony.

The simpler forms of valve transmitter have already been dealt with in Volume I. This chapter is concerned with the application of these principles to medium- and high-powered transmitters of to-day. Except in the simplest forms of transmitter, the valve supplying the power to the aerial is *not* the valve used to generate the oscillations in the first place, but the conditions of operation are largely the same, the fundamental considerations being high efficiency and maximum power output. In the following treatment the valve is considered as an amplifier, the grid excitation being supplied either from a preceding circuit or by reaction coupling from the anode.

**High-efficiency Working.**

In any valve generator, power is supplied to the anode at a relatively high d.c. voltage, and the aim is to convert as much of this power as possible into alternating energy. If the valve is working in the class A condition, i.e. in a condition suitable for distortionless amplification, the maximum theoretical efficiency is 50 per cent. The anode current can vary nearly between the limits of zero and twice the normal steady value, as shown in Fig. 1. Since the r.m.s. value of the alternating component of the current is $I_{max}/\sqrt{2}$, and that of the voltage is correspondingly $V_{max}/\sqrt{2}$, the alternating power is $VI/2$, which is one-half of the steady power consumed at the anode.

In practice, it is not possible to reduce the anode voltage or current fully to zero, and therefore the efficiency is always less than the theoretical amount. It is not necessary, however, to operate the valve in a distortionless condition. All that is required is that a pulse of current shall be supplied by the anode circuit at each oscillation, sufficient

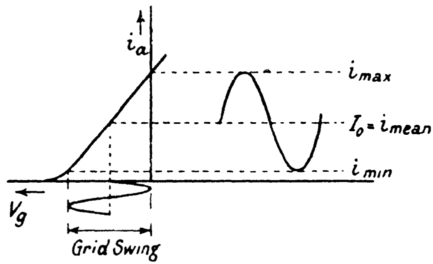



FIG. 1. VOLTAGE AND CURRENT RELATIONS WITH CLASS A OPERATION

to maintain the oscillations. The valve, therefore, is operated with a much higher negative bias than usual, so that the anode current is zero over a large part of the cycle. At certain periods, however, the grid voltage falls to zero and may even become appreciably positive. During these portions of the cycle a large anode current flows and supplies the necessary pulse. The mean anode current, however, is much less than before, and the efficiency of the operation considerably higher.

**Class B Operation.**

Suppose we increase the bias to a value which reduces the anode current nearly to zero, as shown in Fig. 2. Over the positive half-cycle of grid swing the anode current (and hence the anode voltage) will vary proportionally if the characteristic is reasonably straight. Over the negative half-cycle no anode current will flow, but if the anode circuit is tuned the resonant action will maintain the oscillation so that the anode voltage will continue to vary sinusoidally.

Thus, although anode current only flows every half-cycle,

the anode voltage will be a faithful copy of the grid voltage and will alter in magnitude according to the value of the applied grid voltage or excitation. Hence, linear (undistorted) amplification is obtained.

Let $I$ be the steady anode current in the class A condi-

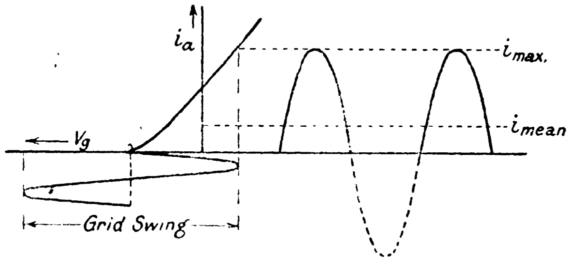



FIG. 2. WITH CLASS B OPERATION THE VALVE IS BIASED TO CUT-OFF

tion. The anode current with class B working is in the form of a series of half sine waves. The average value of a sine wave is $2/\pi$ times the maximum value, but since we are only using every alternate half wave the average value

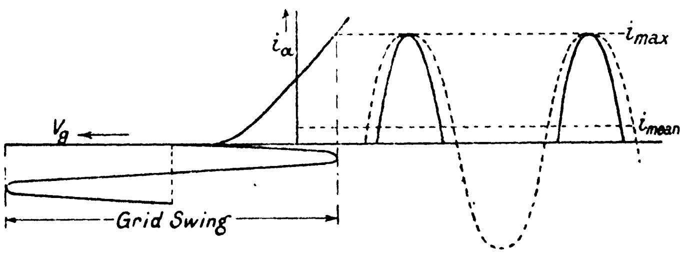



FIG. 3. CLASS C OPERATING CONDITIONS

of the current in this case is $1/\pi$. The maximum value is $2I$, so that the average anode current is $2I/\pi$, which is roughly two-thirds of the class A current. Hence, if the efficiency is 50 per cent in the class A condition, it would be $(\pi/2) \times 50 = 78{\cdot}5$ per cent with class B working. Again, in practice, this is not fully realized, but efficiencies between 50 and 60 per cent are obtainable.

### Class C Working.

Still better efficiencies are obtainable by increasing the grid bias beyond the cut-off point. Under these conditions, sometimes called *flick impulsing,* anode current flows in short pulses of a high peak value, and efficiencies of 70 to 80 per cent are obtained. The anode voltage is no longer directly proportional to the grid voltage, so that the circuit can only be used for maintaining or amplifying a steady oscillation. Where a modulated wave is being handled class B working must be used.

### Matching the Anode Impedance.

The adjustment of a class B or class C transmitter depends upon the matching of the external impedance in the anode circuit to the valve, but the conditions are appreciably different from those of a class A circuit. In general, the optimum impedance is lower than that for maximum power output under class A conditions, which require a load impedance approximately equal to that of the valve.

Fig. 4 illustrates a typical series of characteristics of a transmitting valve. Since we are concerned with obtaining the largest possible anode current, we do not limit the operation of the valve to the negative side of the zero voltage axis, but allow the grid to run positive. It will be seen that after a certain value of grid voltage is reached, little further increase in anode current results, and the thick line on the characteristic is known as the *limiting edge.*

We can draw a load line from the operating voltage $V$ at such an angle that it corresponds with the load in the anode circuit. The line shown, for instance, represents 2 000 ohms, since a change of 400 volts corresponds to a change of 200 milliamperes. The point where this load line cuts the limiting edge represents the maximum peak-anode swing.

Now it will be clear that the lower we make the load (and hence the more vertical the load line) the greater the peak anode current obtained. There is, however, a limit to this, determined by the maximum safe emission of the valve. This is considerably higher than the safe steady

anode current, for two reasons. Firstly, this peak anode current occurs at the point when the anode voltage is at its lowest, and, secondly, it is only momentary. Nevertheless, a limit, known as the *emission limit*, does exist, and our load must be such that it cuts the limiting edge of the characteristic at or below the emission limit.

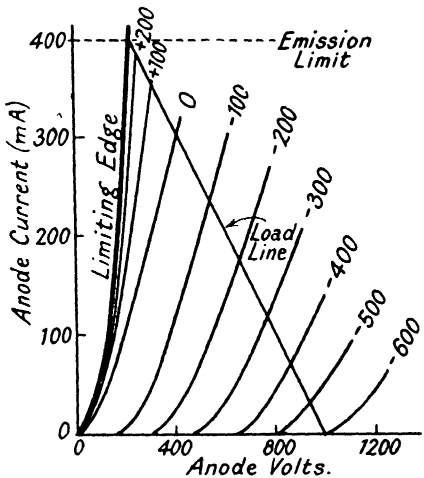



FIG. 4. CHARACTERISTICS OF TYPICAL LOW-POWER TRANSMITTING VALVE

It will be noted that the anode voltage is still not reduced to zero, and, in general, the anode voltage swing is about 80 per cent of the h.t. voltage—a useful empirical rule to remember.

The average anode current depends upon the design of the transmitting valve. Although the maximum anode current occurs at periods of low anode voltage, there is nevertheless an appreciable quantity of power to be dissipated at the anode. This energy is dispersed in the form of heat, and the methods adopted to radiate this heat effectively are discussed later in the chapter. Transmitting valves are usually designed to withstand a peak emission of six to ten times the average value and to dissipate about

25 per cent of the total power. This corresponds to an efficiency of 75 per cent, three-quarters of the total power being converted into alternating-current energy.

Theoretically, the efficiency with class C operation can be made higher than this, but the total power output is found to fall off if this is done. In practice, a compromise is adopted, so that the maximum power output can be obtained.

### Self Bias.

Operation of the valve under these conditions necessitates two adjustments, namely the steady grid voltage and the anode load. The grid voltage may be adjusted by a permanent bias or by allowing the valve to set itself. This is done by including a condenser and leak in the grid circuit. Positive excursions of grid voltage draw grid current from the filament, which charges up the condenser negatively. This process continues until the grid bias is the optimum. Any further increase in negative bias then causes the oscillations to decrease in amplitude so that the grid swing is reduced. The pulse of grid current maintaining the grid condenser charge is no longer sufficient and the grid bias falls again.

The value of grid leak is much lower than in the usual grid detector. Suppose the grid bias became so negative that the valve ceased to maintain the oscillations. The oscillation would not cease immediately but would die away at a rate depending upon the decrement of the circuit. If the charge on the grid condenser can leak away sufficiently rapidly, the grid bias will fall and the oscillation will be maintained; but if the grid leak is too high the oscillation will definitely cease until such time as the grid voltage has fallen sufficiently.

Under these conditions, the oscillation will be continually starting and stopping, giving rise to what is known as "grid tick" or "squegging." The frequency with which the oscillations stop and start may vary from one to every few seconds to several hundreds a second. The remedy is to reduce the value of leak as already explained. About 10 000 to 50 000 ohms is usual.

## Anode Tap.

The effective resistance of a parallel tuned circuit is $L/CR$, where $L$ and $C$ are the inductance and capacitance in microhenries and microfarads, and $R$ is the *total* resistance of the circuit. The valve is adjusted so that this value is equal to the load resistance for optimum working. Alternatively, since this is not always practicable, arrangements are made to tap the anode across a suitable portion of the coil.

For medium and long wavelengths, this tap is usually within the oscillating circuit, but for short waves the reverse may be the case, i.e. the anode may be tapped to an extension of the coil as shown in Fig. 5.

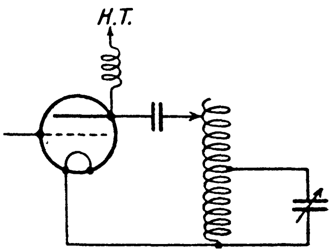



FIG. 5. AUTO-TRANSFORMER CIRCUIT FOR MATCHING THE ANODE LOAD TO THE VALVE

The impedance of a tapped tuned circuit can be calculated by ordinary transformer theory (see page 54 for a typical example), but as a rough approximation we can write $R' = L_1{}^2\omega^2/R$ where $L_1$ is the inductance of the tapped portion and $\omega^2 = 1/LC$.

If $t$ is the tapping ratio $= \sqrt{(L/L_1)}$ (which is roughly equal to $N/N_1$ where $N$ is the total turns and $N_1$ the tapped turns), we can write $R' = (L/CR) \cdot (1/t^2)$ approximately.

In practice, a rough calculation only is made, the actual optimum point being found by trial.

## Energy Transfer.

The resistance $R$ in these formulae is the total resistance in the circuit, including the load resistance. The object of the transmitter is to develop power in some suitable load, usually an aerial, which absorbs power in radiation and is therefore said to have a certain radiation resistance.

From the ordinary laws of coupled circuits (see Chapter IX), the total resistance is $R_1 + (M^2\omega^2/Z_2{}^2)R_2$, where $M$ is the mutual inductance between primary and secondary.

Since the secondary is tuned, $Z_2 = R_2$, so that the expression reduces to $R_1 + M^2\omega^2/R_2$.

To transfer the greatest possible energy to the aerial, the mutual inductance should be as high as possible and the resistance of the closed circuit $R$ should be as low as possible. There is a limit to the increase in $M$, because a coupled circuit has two modes of oscillation (see Chapter IX). If these two frequencies are substantially different the operation becomes unstable and the circuit may jump from one frequency to the other indiscriminately. To avoid this the mutual inductance must be less than $(R_2/\omega L_2)\sqrt{(L_1L_2)}$,

where $L_1$ is the closed circuit inductance
$L_2$ is the total aerial inductance
and $R_2$ is the aerial resistance (including radiation resistance).

In actual practice, this limits the power which can be transferred from primary to secondary to about 70 per cent of the total power, which is one of the advantages in favour of drive circuits. (See Chapter IX for proof.)

**Drive Circuits.**

Instead of maintaining the oscillations by feeding energy back from the anode to the grid of the valve, we can supply energy at the right frequency and phase from an external source. Such a procedure has several advantages. In the first place, more energy can be transferred from the closed circuit to the aerial because the frequency of the oscillation is not controlled by the aerial circuit, and hence the double-hump tuning is no disadvantage. In fact, it may even be of service in maintaining a broad top to the resonance curve, as in a receiving circuit, and the only limit to the energy transfer is the loss in the primary itself.

Secondly, variations in the constants of the aerial and associated circuits do not affect the frequency of the oscillations, but only have a minor effect on its amplitude. This is most important, particularly with short waves. A third advantage is that the modulation, or even the keying of the signal, can be carried out at a stage of the proceedings where the power to be dealt with is only small.

The operation of the valve under driven conditions is practically the same as already described. Such valves are always used as class B or C amplifiers, but it is customary to bias them permanently instead of adopting self-bias arrangements, using grid condenser and leak. This is because a sudden cessation of supply to the grid would automatically release the grid-bias voltage, and the anode

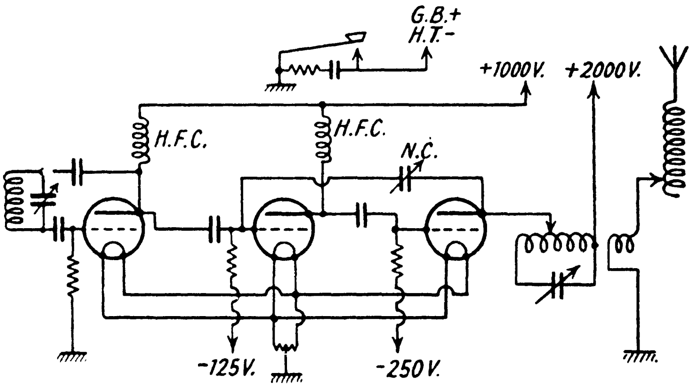



FIG. 6. CIRCUIT OF LOW-POWER TELEGRAPHY TRANSMITTER

current would run up to dangerously high limits, with consequent damage to the valve.

The usual procedure is to arrange a small oscillating valve handling a few watts only. This is followed by an amplifying stage, at which point the modulation is usually introduced. The modulated radio frequency is then further amplified by power stages until the required level is attained, when the signals are applied to the aerial.

As a rule, the first stage is not driven into grid current in order to keep the load on the oscillator small, since this helps to maintain the frequency more constant.

Fig. 6 shows a low-power transmitter utilizing these principles. Keying is obtained by breaking the earthy side of the h.t. supply with a spark suppressor across the key contacts.

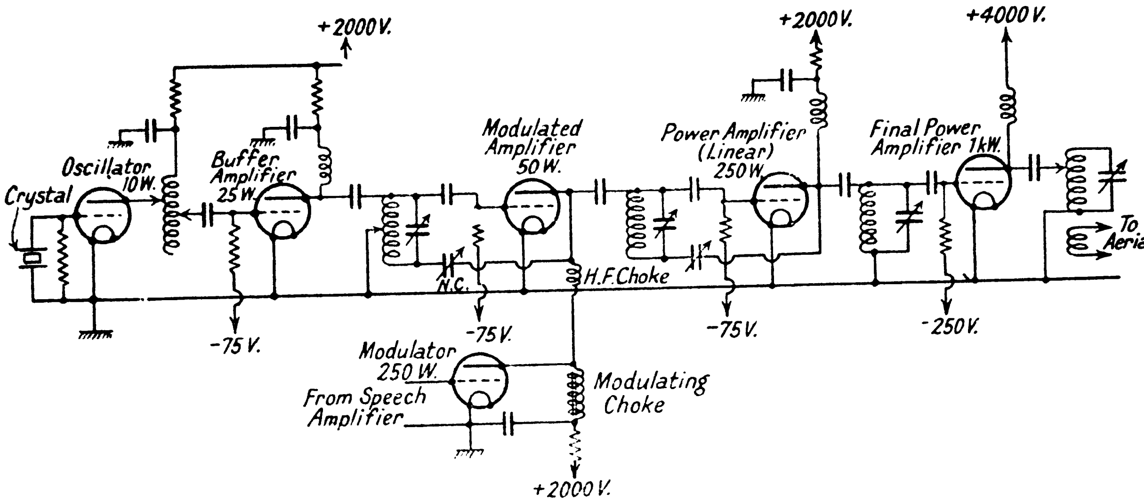



FIG. 7. TYPICAL MEDIUM-POWER TELEPHONY TRANSMITTER

Screen-grid valves are often used in the early stages to avoid neutralizing

## Telephony Transmitters.

In a telephone transmitter the power stages following the modulation cannot be class C operated because such a system is not linear, and it is essential that the amplitude of the oscillations in the anode circuit shall be strictly proportional to the excitation applied to the grid. Class A or B amplifiers must be used, and the second type is usually employed in order to keep the power consumption down.

Fig. 7 shows a medium power telephony circuit. A crystal oscillator (see page 213) is followed by a buffer amplifier which feeds the modulator stage. The modulated signal is then passed through a further amplifier which drives the output valve.

Fig. 8 shows a simplified circuit of a 50 kW. transmitter. A crystal oscillator feeds a buffer amplifier and then a 50 W. amplifier on which the modulation is provided. The modulated output is then amplified by two neutralized 250 W. valves, which then feed two water-cooled valves and then finally feed six water-cooled valves in parallel push-pull.

The power supply is of interest in that three-phase working is adopted. The three phases, being spaced 120° apart, provide a much more uniform output than single phase and require less smoothing in consequence. Single-wave rectification is adopted on the low-voltage supply and double-wave on the high voltage, this providing current pulses every 60°, giving improvement in efficiency. (See Chapter XVII.)

## Screened Valves.

The power amplifiers shown utilize triodes, neutralized to avoid any tendency to self-oscillation, but modern practice is tending to use screen-grid valves. The characteristics of a screen-grid power amplifier intended for positive drive are shown in Fig. 9. The limiting effect with positive grid volts is not apparent here, and while there is the same emission limit it is often not possible to reach this because the load line must be so chosen that equal increments of grid voltage produce practically equal changes of anode current. This is nearly the case in Fig. 9, except at the

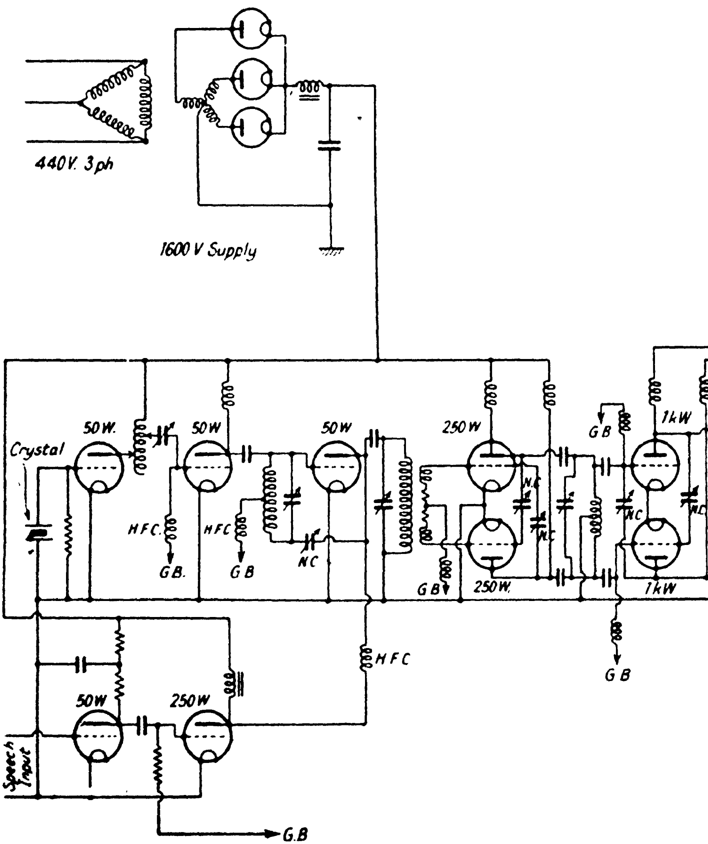
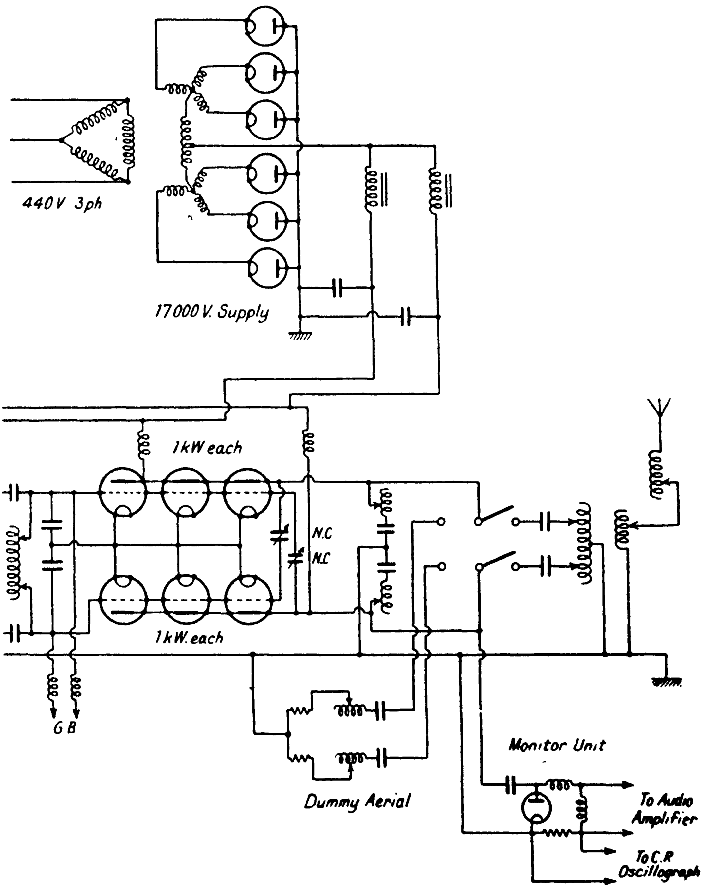



FIG. 8. SIMPLIFIED CIRCUIT OF




A 50 kW. Transmitter

extreme negative end, where some distortion is permissible, since full modulation is rarely employed, and hence the oscillating current never falls to zero. It will be noted that the voltage variation is also limited, owing to the negative resistance effect which occurs if the anode voltage is reduced below the screen voltage as shown dotted on the $Vg = 0$ curve; and, although an improvement is possible with a pentode, screened valves in general give less output than equivalent triodes.

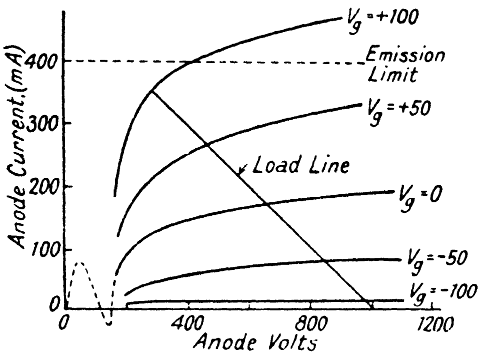



FIG. 9. CHARACTERISTICS OF POSITIVE-DRIVE SCREEN-GRID VALVE

**Modulation.**

Choke modulation is adopted in the majority of instances. The anode supply to the modulating amplifier is obtained through a low-frequency choke, which also supplies the anode of the modulator valve. Variations of the grid voltage of the modulator cause variations in the anode voltage, and since the anode of the modulating amplifier is tied to the anode of the modulator, the voltage varies in the same manner and so alters or modulates the amplitude of the high-frequency oscillations as required.

Fig. 10 illustrates a modulated wave, and it will be seen that with 100 per cent modulation the instantaneous maximum value of the oscillations is twice the mean value of the carrier. Hence, the instantaneous power radiated is four times the carrier power. This has to be supplied in the

form of alternating-current power by the modulator valve, which must therefore be considerably larger than that of the oscillator valve itself.

This discrepancy is to some extent offset by the fact that the mean power developed by the oscillator is not as high as it could be, since the oscillator must be operated in such a condition that it can develop four times its power if required. Under normal conditions of modulation the average power in the modulated condition is about 50 per

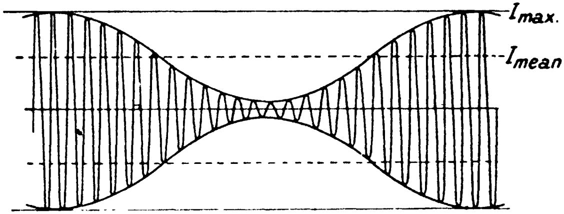



FIG. 10. DIAGRAM OF MODULATED WAVE

cent greater than the carrier power, so that the unmodulated oscillator is only working at some 60 per cent of its full efficiency. Even so, from two to five modulator valves are usually necessary to modulate one oscillator valve of the same type.

At first sight it would seem that if the modulator valve had the same standing anode current as the oscillator valve, conditions would be satisfied, but a little thought shows that this is not so. In the first place, the modulator valve has to work under conditions of undistorted output, which is not an efficient condition, and the anode current variation for full input on the grid is rarely more than 50 per cent of the steady value. Consequently, in order to obtain the necessary full anode current change, at least two, and probably three, valves must be used in parallel or, alternatively, a larger valve employed.

The oscillator must be operated under such conditions that the anode current *and* the oscillating current are both directly proportional to the anode voltage. A class B amplifier or oscillator under suitable conditions will fulfil

these requirements, but it is equally possible that it will not do so and this point must be verified.

**Frequency Modulation.**

Since the frequency of the oscillations generated by a valve is dependent to some extent on the constants of the valve, and since these tend to vary with the anode voltage, it follows that an ordinary self-oscillating valve modulated as described will vary in frequency as well as amplitude during the process. This is undesirable, particularly on the higher frequencies, where a small percentage variation in the frequency may produce a larger change in the received signal than the whole of the modulation.

Modulation, therefore, is seldom employed to-day on the actual oscillating valve, even when special frequency-stabilized circuits are used. The oscillator is isolated by being passed through a buffer amplifier before reaching the modulation circuit, as already shown in Figs. 7 and 8.

**Other Methods of Modulation.**

Other circuits for modulating radio-frequency currents have been devised but are rarely used. The grid-modulated class B amplifier is worth mentioning, as this is sometimes employed for short-wave working. Here the grid bias is caused to vary in accordance with the modulation, and this alters the amplitude of the signals in the plate circuit. This method consumes no energy to speak of and does not require large modulating valves, but it gives greater distortion than choke modulation, particularly at higher values of modulation.

A modification of the choke method is that shown in Fig. 11. Here the h.t. supply to the oscillator valve is obtained through a modulator valve in series. The general action is similar to that of the choke-modulated case, but the system has the advantage that the frequency characteristic is particularly good, i.e. the modulation frequency is immaterial within wide limits. With the choke method the effectiveness of the modulation depends upon the choke presenting a uniform impedance over the whole range of modulation frequency.

The direct modulation method, however, introduces some distortion at high modulation depths. The expression "depth of modulation" refers to the relative percentage of the modulation voltage and the carrier voltage. The expression for a modulated wave is $e = E \sin \omega t(1 + m \sin pt)$

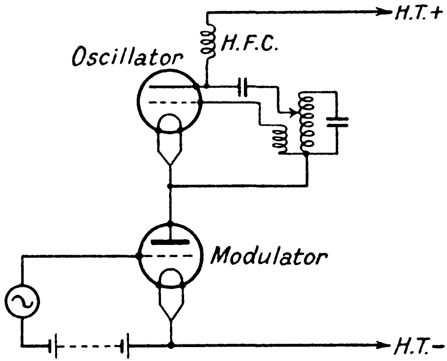



FIG. 11. SERIES MODULATION CIRCUIT

where $m$ is the depth of modulation and $p$ the modulation (angular) frequency. If $m$ is unity we have 100 per cent modulation, since the maximum amplitude of the modulating wave is equal to that of the carrier.

Methods of estimating the modulation depths are dealt with in Chapter XIV.

## Frequency Stability.

The variation of oscillation frequency with the conditions of the circuit and the valve have already been mentioned. It is important to minimize these variations as much as possible. The coils and condensers of the fundamental oscillating circuit should be so constructed as to vary little with temperature and other atmospheric conditions. This is an added argument in favour of using a small oscillator in the first place, since it is obviously easier to adopt such a construction when the dimensions are small.

The next requirement is to minimize the variations caused by changes in the valve. The grid-filament impedance of

the valve is shunted across one portion of the tuned circuit and the anode-filament impedance across another portion. Any variation in these impedances will affect the frequency slightly, for the resonant frequency of a tuned circuit is not entirely controlled by the inductance and capacity,

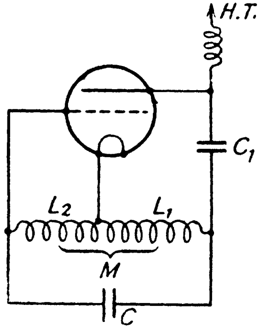



FIG. 12. TYPICAL CONSTANT-FREQUENCY OSCILLATOR CIRCUIT

$$C_1 = \frac{L_1 + L_2 + 2M}{L_1 + L_2 A^2 - 2MA} C \text{ where } A = \frac{L_1 + M}{L_2 + M}$$

but is also affected to a small extent by the resistance in the circuit. The presence of the valve resistances across certain parts of the circuit alters its effective resistance and hence affects the frequency.

It can be shown that stable operation is obtained if either the anode resistance or the grid resistance of the valve can be maintained constant. One method of maintaining stability, therefore, is to use the highest value of grid leak which can be employed without intermittent operation (ticking). Under these conditions, the grid resistance of the valve is approximately half that of the grid leak, and therefore tends to remain constant.

An alternative method is to include a resistance or a reactance in series with the anode of the valve. The effect of variations in the valve resistance is thereby minimized.

Thirdly, a circuit with a high $Q$* is an advantage,

* The symbol $Q$ is used for the magnification of a circuit $= L\omega/R$.

because a change of frequency then produces a rapid change of phase, and this exercises a correcting influence on the frequency.

## Constant Frequency Oscillators.

It is impossible to discuss here the theory underlying this statement, but by following up this question of phase displacement it is possible to devise oscillators in which the frequency is independent of variations in the valve characteristics. The reader is referred to a paper entitled "Constant Frequency Oscillators," by F. B. Llewellyn, in the *Proceedings of the I.R.E.*, December, 1931. The procedure, in effect, lies in introducing into either grid or anode lead a phasing impedance of such a value that any change of phase due to frequency variation alters the frequency by the exact amount required to keep the frequency constant. Fig. 12 shows a typical circuit.

## Electro-mechanical Oscillators.

The increasing call for extreme stability of frequency, necessitated by modern conditions, has led to the development of methods embodying mechanical resonant circuits. When a large number of channels is required, operating within a limited band of frequency, the band width per channel has to be strictly adhered to. Any serious wandering of the carrier frequency will cause the side bands to overlap, producing an unintelligible but noticeable "cross-talk," known as *side-band splash*.

A mechanical resonator is very sharply tuned and can be kept operating within limits of a few parts in a million. One commonly used device is the quartz crystal. A suitably cut plate of quartz, subjected to electrical stress, will oscillate longitudinally (i.e. by contraction and expansion about one of its axes), and this forms the basis of a very stable oscillator. The subject is discussed more fully in Chapter XII.

For long-wave transmissions, use has also been made of tuning forks and magneto-striction oscillators. A tuning fork will only operate at an audio frequency, and hence

several stages of frequency multiplication must be used to arrive at the required radio frequency.

## Magneto-striction.

Magneto-striction oscillators are sometimes used, while the effect has also been employed in other directions. Some materials, such as stainless steel, cobalt steel, nickel and invar, show an appreciable change in length when subjected to a magnetic field. A rod of the material, housed in a long solenoid carrying a.c., will thus vibrate longitudinally at the frequency of the a.c. (provided a d.c. polarizing current is also applied to prevent the magnetic field from reversing. Otherwise it will vibrate at twice the applied frequency, since the magneto-striction effect is independent of the direction of the magnetism).

An oscillator can be constructed by coupling two coils to the rod, each extending from near the centre to the end, and connecting these in the anode and grid circuit of a valve. A small current in the anode coil will set the rod vibrating, and this will induce voltages in the grid coil, which, if correctly phased, will provide amplified currents in the anode coil and hence build up the oscillation. The rod will vibrate as a half-wave resonator, so that the centre point is a node and may be rigidly clamped.

The frequency of vibration is determined almost entirely by the rod, irrespective of the circuit conditions, and is given by $v/l$, where $v$ is the velocity of sound in the material and $l$ the length of the rod. Frequencies ranging from one to several hundred kc/s. can be produced by this means, the upper limit occurring when the skin effect in the material prevents the magnetic field from penetrating to an adequate depth. The stability is far better than ordinary methods, but not equal to that obtainable with good quartz oscillators.

For further information, the reader should refer to a paper by G. W. Pierce, *Proc. I.R.E.*, Jan., 1929.

## Coil Construction.

The heavy currents to be carried in a high-powered transmitter, particularly in the later stages, necessitate

special conductors. Solid wire is not used in general practice owing to the skin effect, and copper tube or strip is usually employed. Loosely-woven copper braid is employed for flexible connexions. Stranded (Litzendraht) cable is very expensive in large sizes and is used sparingly.

The heating of the coil causes expansion of the turns, which is a factor requiring attention during the design of the coil. Any simple wind on a former is unsatisfactory, because as soon as the coil heats up in use the turns will become slack, and with alternate heating and cooling the mechanical rigidity of the coil is very soon destroyed.

A built-up form of construction is usually adopted, partly to allow for this and partly to minimize the dielectric losses as explained in the next section. The former is made with a series of longitudinal ribs on which the wire, tube, or strip is wound. The turns are located in grooves and are rigidly held at these points. The number of ribs must be so chosen in relation to the thickness of the wire and the diameter of the coil, so that under the worst conditions of overheating likely to be experienced in practice, the wire will not become slack enough to affect the operation, and any deformation will not be permanent.

## Dielectric Materials.

The material for the insulation also requires careful selection. Eddy currents circulate in the dielectric and generate heat, which not only absorbs power from the circuit, but may constitute an actual fire risk if the material is not suitable. Hard rubber compounds, such as ebonite, are not usually employed except at quite low frequencies, for the sulphur content of the material overheats due to the influence of the radio-frequency fields, and with short waves this effect is so marked as to render the use of ebonite quite impossible.

Paper products impregnated with suitable composition are widely used. Thin sheets of paper are either compressed together to form one thick slab, or are rolled up to form rod of circular or other section. The paper is impregnated with bakelite varnish or other special insulating

varnish, and in some cases linen is employed to give greater mechanical strength. This class of product goes by various trade names, such as Paxolin or Tufnol.

Mica is considerably used for small parts, but its brittleness seriously limits its application. There are, however, various mica products formed by bonding powdered mica with suitable materials and forming sheets or slabs. A particular example is a synthetic red coloured material composed of rubber and mica, known by various trade names such as Keramot, Silvonite, etc. It is very good for medium radio frequencies, and combines low losses with high dielectric strength (55 kV per mm.).

A material which is much used, particularly for high-frequency work, is Mycalex, which is ground mica bonded with lead borate glass and moulded under pressure at about 675° C. It is a dull grey material which can be machined with special tools. A number of other synthetic materials are coming to the fore.

Porcelain is suitable for many applications. This is a ceramic material containing mainly china clay (kaolin), flint, and feldspar, which are finely ground and mixed, and then mixed with a small quantity of moisture to enable them to be moulded. They are then fired and, for some applications, glazed to provide a smooth surface. It is also possible to use a dry casting process in which the powder is formed into the required shape under pressure without firing.

Similar ceramic compounds have been highly developed, particularly for short-wave work. Magnesium silicate is also used instead of clay (which is an aluminium silicate); while steatite, which is a purified talc sometimes known as soapstone, is also utilized.

The ceramic insulators are probably the best from the point of view of dielectric properties, but because of their fragility they are only used in positions where the highest efficiency is essential.

## Layout.

The transmitting apparatus itself is usually housed in one or more cubicles of steel frame construction, so arranged

that access to the "live" parts cannot be obtained without disconnecting the supply.

This is an important proviso in view of the high voltages involved, and considerable care has to be taken to make the operation safe.

In many instances, as in the transmitters of the B.B.C., for example, each stage in the chain is duplicated, including the valves, and by throwing over a switch the current may be routed through the alternative stage. This enables continuity of service to be maintained. Both channels are tuned up and the valves kept alight, so that a simple change over of connections is all that is required.

The main controls are on the transmitter, but controls of input, modulation, etc., are effected from a remote control desk, on which are duplicated the principal circuit meters and a monitor of the output, so that the control engineer can see the situation at a glance.

## Valves for Transmitters.

Transmitting valves are similar to those employed for receiving purposes, except that they are larger in size and are designed to dissipate heat very readily from the anode and to a smaller extent from the grid. The valves are designed to give a high peak emission, the ratio of peak emission to mean anode current being of the order of 10 to 1.

In order to increase the peak dissipation at the anode, materials are used which will run at a higher temperature. Molybdenum is one such material, and anodes made with this will run at a cherry red. Another method of increasing radiation is to blacken the anode, since a black body radiates heat better than a bright one. Still another method is to use carbon for the anode, and this is increasingly employed to-day. The main difficulty is to get rid of the occluded gases, but with improving methods of technique this is becoming easier.

Heat dissipation is by direct radiation from the anode to the glass and thence ordinary convection. Since a large bulb has to be employed, in order that the glass shall

FIG. 13. X-RAY PHOTOGRAPH OF MARCONI HIGH-POWER WATER-COOLED TRANSMITTING VALVE
*(By courtesy of Marconi's W. T. Co.)*

not soften, this ultimately sets a limit to the size of the valve. Forced draught cooling has been employed, and even oil-immersed valves have been used, but the improvement was not sufficient.

The next step was in the direction of silica valves, since this material will run at a higher temperature than glass. A silica envelope, however, is very expensive, and valves of this type are usually made demountable, the construction being such that if the filament burns out the base of the valve can be removed, a new filament fitted, the base relined up and welded to the envelope, and the whole re-exhausted.

The latest and undoubtedly the most successful technique is that of using cooled-anode valves. Here the anode itself is made the outer container of the valve. The grid and filament assemblies are mounted on a glass foot as usual, and this is sealed into the anode. For medium powers this is sufficient in itself. The anode is usually formed with large cooling fins to increase the radiation.

For still larger powers the anode itself is immersed in a water jacket, around which a constant supply of water circulates. This keeps the anode cool and dissipates the heat very satisfactorily.

Since the anode is at a high potential, the water circuit has to be suitably insulated, and this

is usually done by connecting the supply to the anode through a long length of rubber hose suitably coiled up. Valves of this type are in daily use handling powers of 25 and 50 kilowatts each, and banks of such valves handling radio-frequency outputs of 500 kilowatts or more are employed in modern technique.

Certain specialized developments are necessitated by short-wave working, but these are discussed more fully in Chapter XII.

## EXAMPLES I

(1) In a transmitter operating on 377 metres the closed circuit inductance is 200$\mu$H. and the resistance 5 ohms. If the aerial inductance is 50$\mu$H. and the aerial resistance is 5 ohms, calculate

(a) The maximum mutual inductance permissible;

(b) The percentage of power transferred to the aerial. (Estimate this in terms of the increase in primary resistance due to the aerial.)

(2) If double-humping is to be avoided, the volt-amperes in the secondary must be less than that in the primary. On this basis prove that the limiting coupling is

$$M = (R_2/\omega L_2)\sqrt{(L_1 L_2)}$$

(The volt-amperes are the product of the volts across the coil (or condenser) and the current in the circuit.)

(3) Calculate the approximate tap on the closed circuit of question (1) to match the circuit to a valve having an optimum load of 1 600 ohms.

# CHAPTER II

## THE TRANSMITTING AERIAL

HAVING generated radio-frequency oscillations in a radio transmitter, it is then necessary to cause these oscillations to radiate wireless waves as effectively as possible. To do this it is necessary to cause the current to flow backwards and forwards in an elongated wire known as the *aerial* or *antenna*.

The idea of a wave being produced as a ripple in a line of force, rather like a ripple produced by jerking a rope with the hand, has already been propounded in Volume I, and it will be clear that the larger we can make this ripple the stronger will be the wave produced.

Our aim, therefore, is to cause the electric current to flow between two points as far apart as possible, and for most purposes of radio-frequency transmission these points are situated vertically apart. Hence, it becomes necessary to erect a mast or tower of some kind to support a vertical wire into which the current may be fed. The passage of the electrons up and down this wire will produce the desired radiation.

### Natural Wavelength.

Now, since a wire possesses inductance and capacitance of its own, there will obviously be some particular frequency at which the inductance and capacitance of the wire resonate so that the aerial itself constitutes an oscillating circuit. This occurs when the length of the wire is about 80 per cent of the wavelength of the radiation which would be generated by current of that particular frequency. An aerial in such a condition is said to be operating at its *natural wavelength*, and it is interesting to consider such an aerial in detail.

The inductance and capacitance are not concentrated, but are more or less uniformly distributed over the whole length.

Consequently, the current and voltage will not be uniform, but will vary from place to place along the wire.

The current at the end of the wire must clearly be zero, because there is no other path for the current, and since the arrangement is symmetrical the current at the centre will also be zero, the distribution being sinusoidal in form as shown in Fig. 14 (*a*).

The current is thus in the opposite direction in the two

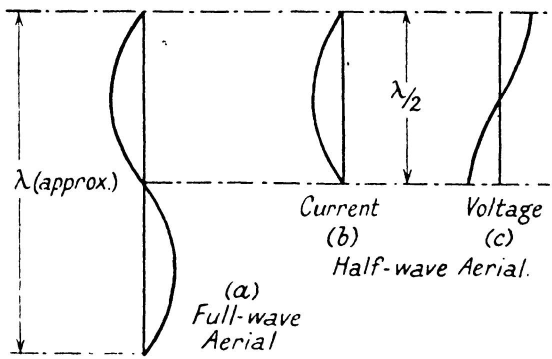



FIG. 14. DISTRIBUTION OF CURRENT AND VOLTAGE ON A WIRE OSCILLATING NATURALLY

halves of the aerial, so that the radiations would cancel out. If we want to radiate we must limit the length to a half wavelength as shown in Fig. 14 (*b*), where we have zero current at each end and a maximum in the middle.

The instantaneous value of the current at the centre (or at any other part) will, of course, be varying from instant to instant according to the oscillation frequency, but if we consider the r.m.s. value, we shall find that the distribution of current is of the form shown.

Now, in an oscillating circuit we know that when the current is zero the voltage is a maximum, and vice versa, and hence the voltage distribution will be as shown in Fig. 14 (*c*), with the maximum voltage at each end and zero voltage in the middle. The voltages at the two ends will of course be equal and opposite.

**Earthed Aerials.**

If, instead of using a half-wave aerial such as this, freely suspended in space, we connect the bottom end to the ground we find that the current in the aerial induces currents in the earth, which produce the effect of an image of the aerial in the ground. This enables us to reduce the aerial length by half, so that we have a quarter wavelength above the ground, the remainder of the aerial being supplied by the image in the earth itself.

This clearly gives a convenient arrangement, for it brings

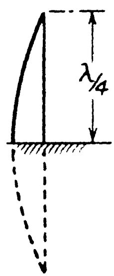



FIG. 15. QUARTER-WAVE AERIAL

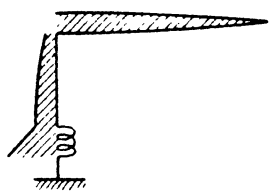


FIG. 16. CURRENT DISTRIBUTION ON LONG-WAVE AERIAL

the point of zero voltage and maximum current actually at the ground, which is what we want. The aerial would be coupled through a coil or by other suitable means to the closed oscillating circuit of the transmitter, and would definitely be connected to earth at the bottom end.

The question now arises as to whether we can use an aerial of this type. The production of a wire a quarter wavelength long is rather difficult when we are dealing with wavelengths of the order of, say, 3 000 metres. The wire would have to be nearly half a mile high, which is impracticable.

We have two alternatives. One is to make the wire as high as we can and then bend over the top for the remainder of the required length. This can be done up to a point, provided the horizontal top is not too long relative to the vertical height. Otherwise the major part of the current is

providing horizontal radiation down into the ground or up into space, which is not what we want. The other alternative is to include a loading inductance at the bottom of the aerial—in other words to coil up the bottom part of the aerial in a confined space.

This second alternative enables us to reduce the actual length of aerial wire required to tune the aerial circuit to the required frequency, but it has the disadvantage that the maximum current is flowing in the coil, which has poor radiating properties, while the current which flows in the aerial itself is only the relatively small current which exists towards the top of the aerial, being actually nothing at the far end. In practice, therefore, a compromise between the two methods is used for medium and long waves, the aerial being provided with a flat top, usually from one to three times the actual height, and a loading inductance included at the base of the aerial in order to bring the aerial up to the tuning point.

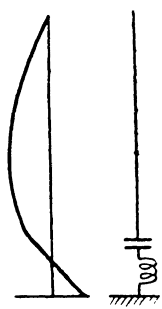


FIG. 17
ILLUSTRATING USE OF "SHORTENING" CONDENSER

As one increases the frequency of the oscillation the aerial length is proportionally reduced, and it becomes practicable to employ aerials half a wavelength long. A discussion of this class of aerial is given in Chapter XII, dealing with short-wave operation. Reference may be made, however, to a type of aerial occasionally employed for the shorter medium waves ranging from 150–250 metres.

In order to maintain a large effective height, masts 200–300 ft. high may be employed, and the length of the aerial in this case becomes greater than the quarter wavelength. It is then no longer practicable to feed in current at the "node" in the centre, nor is the earth point at zero potential. A loading inductance is of no value here, and it is necessary to insert a condenser in the aerial circuit as well as an inductance, and to adjust these so that the total effective length of the aerial is three-quarters of a wavelength. The earth, then, again becomes

a point of maximum current and zero voltage, as shown in Fig. 17.

### Directional Radiation.

Where the aerial is of the same order as, or several times greater than, the wavelength of the transmission in question, the radiation is not by any means uniform in all directions. This point is discussed in more detail in Chapter XII, but it is worth noting that even with aerials appreciably shorter than the natural wavelength, directional properties are obtained. With a flat-topped aerial, for example, the radiation from the horizontal portion interacts with that from the vertical portion and strengthens the radiation in one direction while diminishing it in the other. An "L" aerial, such as that shown in Fig. 18, would thus tend to radiate upwards and more strongly away from the horizontal top, as indicated in the figure.

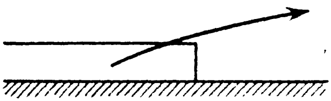


FIG. 18. AN "L" AERIAL IS PARTIALLY DIRECTIONAL

### Non-fading Aerials.

For broadcasting, a special technique is often adopted. In the immediate neighbourhood of a station the transmission is accomplished by the ground wave radiated horizontally from the transmitter. After a short time this dies out and reception is accomplished practically entirely by the indirect ray which is reflected from the Heaviside Layer, as described in Volume I, Chapter XXIX. Transmissions by the indirect ray are liable to fade from time to time due to periodic variations in the Heaviside Layer or rotation of the plane of polarization of the wave during reflection, and therefore the broadcast engineer only considers his station to be effective over the area served by the ground ray.

Unfortunately, before the ground ray has really died out, interference begins to creep in from high-angle radiation which leaves the transmitter in an upward direction

and is sharply reflected by the Heaviside Layer so that it returns to earth in a comparatively short distance (anything from 100 miles upwards with ordinary broadcast wavelengths). This high-angle radiation suffers from fading and distortion in the same manner as the longer range indirect ray and if it is comparable in strength with the ground ray at any given point serious fading will obviously be obtained.

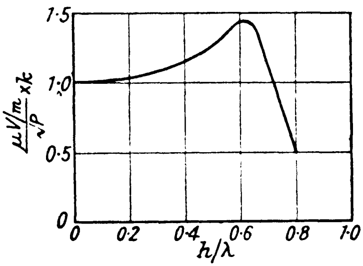



FIG. 19. ILLUSTRATING RELATIVE INCREASE IN HORIZONTAL RADIATION WITH HALF-WAVE AERIAL

There has, therefore, been some research of recent years into special forms of aerial which minimize this high-angle radiation and at the same time accentuate the horizontal radiation, both of which will increase the ratio of ground wave to sky wave at a given distance from the transmitter and consequently will increase the service area of the station. Fortunately, the form of aerial required is relatively simple.

Suppose we consider a simple vertical wire half a wavelength long. Each portion of the wire will radiate energy according to the current at that point. Consider a wave generated by the current at the base of the aerial and travelling in an upward direction. At the same instant a similar wave is being generated by current at the top. But this second wave will be ahead of the wave produced by the base, and in fact by the time the base wave has reached the top, the current at the top will be flowing in the opposite direction and will generate a wave which will cancel out the wave which arrives from the base. This will be clear

because the time taken for a wave to travel half a wavelength is exactly half a period, so that the current will have completely reversed.

Waves leaving from intermediate points on the aerial will of course not completely cancel out, but it will be clear that the upward radiation will be distinctly limited. Moreover, if we make the aerial still more than half a wavelength long the current in the top portion will begin to emit reverse radiation to an increasingly strong extent, for beyond half a wavelength the current reverses as shown on page 27.

On the other hand, the ground wave will suffer if the aerial is made much more than half a wavelength long because the radiations in the horizontal plane are in phase and hence the waves from the top portion will oppose those from the bottom portion due to the reversal of the current. Experiments show that an optimum condition is obtained when the length of the aerial is about $0{\cdot}6\lambda$, and a number of these aerials have been erected recently both in America and Europe.

The usual procedure is to make the mast itself the aerial. It is built in a gradually increasing section up to about $0{\cdot}25\lambda$ after which the section decreases again. The ultimate section is often a simple vertical rod on top of the mast proper. The base of the mast has to be insulated because it is at the maximum potential, and it must be connected to the top of an oscillating circuit equivalent to a quarter-wave aerial to bring the earth point at low potential. It is fed through feeders from the transmitter in the usual way. (See next chapter.)

Fig. 19 shows the improvement in conditions with increasing height. The vertical axis is the horizontal field strength divided by the square root of the power input, which is the criterion which concerns the engineer whose object is to produce the maximum horizontal radiation for a given power. It will be seen that with a height of $0{\cdot}6\lambda$ there is an increase of nearly 50 per cent in the horizontal radiation, while as explained at the beginning the high-angle radiation is limited because of the opposing effect of the radiation from the top portion of the aerial.

Readers who wish to follow the subject further should refer to an article by Ballantine, entitled "High Quality Radio Broadcasting," *Proc. I.R.E.*, May, 1934, Vol. 22, page 616.

The conditions are different when we are dealing with short waves, for here it is possible to use aerials of a length

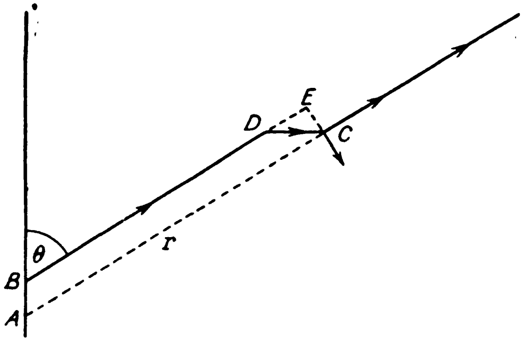



FIG. 20. ILLUSTRATING RADIATION FROM A SHORT LENGTH OF WIRE

comparable with the wavelength, and this demands special treatment, which is considered in more detail in Chapter XII.

## Formation of Waves.

In assessing the actual field strength radiated by an aerial, it is necessary to consider the radiation from a very small portion of the wire in which the current may be assumed constant, and then to integrate the effect of all such small elements with due allowance for the gradually diminishing current as we approach the end of the aerial.

It is instructive to consider briefly the mechanism of radiation.* Fig. 20 represents a line of force from an electron at $A$ moving uniformly upwards with a velocity $v$.

Let us now suppose that for a short time $\delta t$, the electron, accelerates to $B$. Its new velocity will be $v + a \,.\, \delta t$, where $a$ is the acceleration, and the new line of force will radiate

* For a rigid treatment, the reader should refer to *Wireless* by L. B. Turner (Cambridge University Press), the first two chapters of which are very helpful in clarifying fundamental conceptions.

from $B$. But any disturbance is transmitted through the ether at a finite velocity $= c$, the velocity of light. At a distant point, therefore, the information regarding the new position of the electron will only arrive after a definite lapse of time.

Actually after a time $t$, the information will have reached a point $ct$ distant, so that anywhere beyond a sphere of this radius the lines of force still emanate from $A$. Inside a sphere of radius $c(t - \delta t)$ they will emanate from $B$. In between the two the lines are "kinked" as shown, and it is this kink which constitutes the wireless "wave."

We can resolve the "kink" $CD$ radially and tangentially. The electric field along a line of force (from the fundamental laws of electrostatics) is $e/r^2$, where $r$ is the distance from the electron. Along $DC$ it will be

$$\frac{e}{r^2} \Big/ \cos CDE$$

Resolving this radially and tangentially, we have

$$\text{Radial component} = \frac{e}{r^2 \cos CDE} \cdot \cos CDE = \frac{e}{r^2}$$

Tangential component

$$= \frac{e}{r^2 \cos CDE} \sin CDE = \frac{e}{r^2} \cdot \frac{CE}{DE}$$

The radial component is thus the same as the normal radial field, while the tangential component is the radial field multiplied by $CE/DE$.

$$\text{Now, } CE = AB \sin \theta = at \,.\, \delta t \,.\, \sin \theta$$

while if the point $C$ is far from $A$, $DE$ is substantially equal to the difference in radius of the two spheres.

I.e. $$DE = ct - c(t - \delta t) = c\delta t$$

$$\therefore \quad \frac{CE}{DE} = \frac{at \sin \theta}{c} = \frac{ar \sin \theta}{c^2} \text{ since } r = ct.$$

$$\text{Hence the tangential component} = \frac{e\, a \sin \theta}{c^2 r}$$

It should be noted that this expression, denoting the field strength of the wireless wave, is

(*a*) Proportional to the acceleration $a$, so that the more rapidly we accelerate the electrons (i.e. the higher the frequency of the oscillation) the greater the effect.

(*b*) Inversely proportional to the radius $r$ whereas the normal induction is proportional to $1/r^2$ and hence rapidly becomes negligible.

(*c*) Proportional to $\sin \theta$, and is therefore greatest in the horizontal direction and zero in the azimuth.

**Radiation from an Aerial.**

We now have to apply this reasoning to the practical case of an aerial. Consider a current $i = I \sin \omega t$ flowing in a wire of length $l$. Let $v$ be the velocity of the electrons in the wire. Then the current, in terms of the total charge on the electrons, $e$, is $ev/l$ in electrostatic units or $ev/l \times \dfrac{1}{3 \times 10^9}$ amps.

$$\text{Hence } v = 3 \times 10^9 \frac{l}{e} I \sin \omega t$$

$$\text{The acceleration } a = \frac{dv}{dt} = \frac{3 \times 10^9\, \omega l I \cos \omega t}{e}$$

By substitution in the formula previously obtained, we have that the electric field strength

$$\mathbf{e} = \frac{3 \times 10^9 \omega l I \sin \theta}{c^2 r} \cos \omega t\ *$$

This is in electrostatic units and to convert it to volts we multiply by 300, giving

$$\mathbf{e} = \frac{9 \times 10^{11}\ \omega l I \sin \theta}{c^2 r} \cos \omega t \text{ volts/cm.}$$

The simplest form of practical aerial is the Hertzian doublet, consisting of two plates or bodies joined by a wire

* Strictly, the last term should be $\cos \omega(t - r/c)$, but this can be neglected in a simple exposition such as the present.

(Fig. 21). The capacitance is assumed all concentrated in the ends and the inductance of the wire is neglected so that the current is uniform. A practical long-wave aerial is equivalent to half of a doublet, the lower half being formed by an "image" below the ground. Writing $h$ for the height

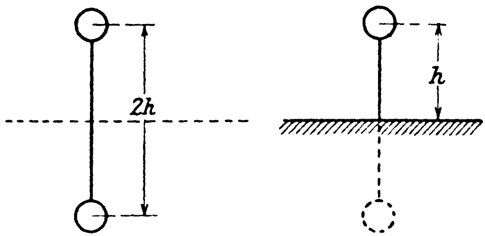



FIG. 21. THE HERTZIAN DOUBLET

of the half doublet, $\omega = 2\pi c/\lambda$ and $c = 3 \times 10^{10}$ cm. sec., our expression for field strength becomes

$$\mathbf{e} = 120\pi \frac{Ih}{\lambda r} \sin\theta \cos\omega t \text{ volts per cm.}$$

For many purposes we are concerned only with the radiation in the horizontal plane, when $\sin\theta$ becomes 1. Then in r.m.s. terms, and writing $h$, $\lambda$ and $r$ in metres, we have

$$\mathbf{e} = 377\, Ih/\lambda r \text{ volts per metre.}$$

These are the fundamental expressions for radiation from an aerial and show that the field strength at any distance from the transmitter is proportional to the aerial current and (effective) height and inversely dependent on the wavelength and the distance.

**Effective Height.**

In practice the effective height is less than the actual height, even with a long-wave aerial, the ratio being of the order of 70 per cent, while if the aerial does not have a large top the formulae are modified by the fact that the current

is no longer uniform but is sinusoidal as shown in Figs. 14 and 15. In such a case the effective height may be taken as the actual height multiplied by the average value of the current, which for a sinusoidal distribution is $2/\pi$.

For other distribution similar allowance must be made. For example, if the aerial is loaded with inductance to several times its natural wavelength, but has no flat top, the current would fall practically uniformly from bottom to top and the effective height would be $h/2$.

## Radiation Resistance.

From the expression for the field strength it is possible, by integrating the values through the whole of the hemisphere surrounding the aerial, to arrive at an expression for the total power radiated, which comes to $160\pi^2h^2I^2/\lambda^2$, where $I$ is the r.m.s. value of the aerial current. If this power is assumed to be dissipated in a mythical "radiation" resistance, we can say that $R_{rad} = 160\pi^2h^2/\lambda^2 = 1584h^2/\lambda^2$.

For a quarter-wave aerial (Fig. 15) $R_{rad} = 160\pi^2h^2/\lambda^2 \times \frac{4}{\pi^2} = 40$ ohms (since $\lambda = 4h$).

## Measurement of Field Strength.

When a new transmitter is to be installed, tests are often made, where practicable, to try and estimate the field strength likely to be received. Later, the radiation from the completed station is checked in the same way and a chart is drawn up to show the service area.

The field strength is measured in microvolts per metre, being the actual strength of the electric field at the receiving point. If we had an aerial of one metre effective height, and this was influenced by a wireless wave having a strength of ten microvolts per metre, then the voltage picked up by the aerial would be ten microvolts.

If the aerial is coupled to a tuned circuit, the voltage developed across the circuit will, of course, be several times greater than this owing to the amplification of the circuit, but even so it is too small to measure as such. Field-strength measurements, therefore, are usually made by

providing an artificial transmitter which is capable of inducing in the aerial a known very small voltage.

The test aerial is then connected to a suitable tuned radio-frequency amplifier, and the output voltage in the detector stage is noted. By means of a changeover switch, the voltage is now induced from the artificial transmitter, which is adjusted to give the same output voltage as before. It is thus possible by comparison to determine the induced voltage in the aerial and hence, knowing the effective height of the aerial, to calculate the field strength in microvolts per metre.

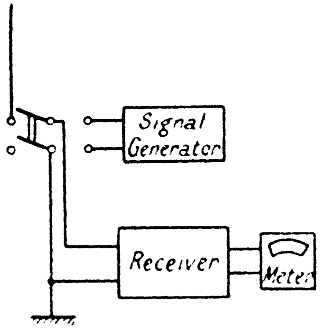


FIG. 22. SET-UP FOR MEASURING FIELD STRENGTH

The input voltage can be estimated in various ways. The usual method is to feed the current from the local oscillator through a calibrated attenuator. From a knowledge of the initial current and the attenuation the voltage in the dummy aerial can be estimated with satisfactory accuracy. It should be noted that the amplification of the receiver, which is of necessity liable to vary from time to time or even with the setting, does not enter into the calculation, but it is, of course, essential that the local oscillator shall be very carefully screened.

The table shown on page 39 indicates the order of field strength experienced in modern practice.

It is worth noting that a good foreign station such as Leipzig or Moscow provides a field strength in England of the order of 1 to 2 millivolts per metre.

## The Earth System.

Efficient radiation with an earthed aerial depends to a large extent upon having a good earth connexion, so that the ground may act as an efficient mirror as already described. For this purpose the transmitting site is usually located on ground which has a good conductivity, and

| Field Strength $\mu$V per m. | Remarks |
|---|---|
| 10 000 (10 mV.) | Very good signal, well audible above local interference. |
| 1 000 (1 mV.) | Good signal requiring a measure of h.f. amplification, liable to interference but capable of good service. |
| 100 | Fair signal requiring good tuning circuits and high amplification. Will be subject to serious interference. |
| 10 | The smallest commercial signal. Will require full resources of selective amplification, directional aerial, and such-like to give service through interference, atmospherics, and local electrical noise. |

the earth connexion is made by burying a series of wires or pipes running from the transmitting building. These may be concentrated immediately under the aerial system or may run in all directions, the deciding factor being largely one of cost.

Earth losses are important from the viewpoint of overall efficiency and attempts have been made to reduce them to negligible proportions by using a *counterpoise* or *earth screen*. This consists of a network of wires a few feet above the earth running under the aerial and for a considerable distance each side.

Experience shows that, just as a network of wires in the aerial system has an efficiency nearly as great as would be obtained by a solid sheet, so an earth screen of this nature constitutes an efficient mirror with less loss than would be obtained from a conventional earth connexion.

In a well-designed aerial system, however, the whole conductor loss (including earth loss) is only some 30 to 40 per cent of the total aerial resistance (see Volume 1, Chapter XXXII), and a point is thus reached where it is not economical to improve the efficiency of the earth system. For this reason the majority of transmitters content themselves with a good buried earth.

# CHAPTER III

## FEEDERS

In a simple transmitting system it is possible to lead the aerial straight into the transmitting building, and to situate the aerial tuning inductance close to the transmitter. This, however, is not always convenient. One may have several transmitters to be housed in the same building, and it is obviously better if the aerial systems can be kept well spaced from one another to avoid any possibility of interaction. Consequently, attempts have been made to carry the energy from the transmitting building to the

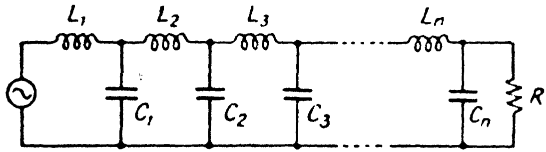



FIG. 23. ILLUSTRATING FEEDER AS A SERIES OF SECTIONS

aerial through a transmission line, in much the same way as power is carried at low frequencies for industrial purposes.

The development of this feeder technique received a considerable impetus following the introduction of short waves, for some type of feeder is essential in the modern short-wave transmitting station. We will, therefore, discuss briefly the technique of the transmission of energy through feeder wires.

Suppose we consider two wires of indefinite length connected at the sending end to a source of a.c. voltage. The wires will each have a small self-inductance which will be uniformly distributed along their lengths. There will also be a capacitance between them, likewise uniformly distributed along the wires. Let us break the feeder up into small sections and consider the inductance and capacitance concentrated in each section, so that the system looks like Fig. 23.

The voltage at the sending end will cause a current to flow into the first condenser. This current will lag slightly behind the voltage, because of the inductance in the wires. The condenser will build up a charge, again with a slight time lag, so that voltage will appear across $C_1$, similar to the input voltage but slightly delayed.

This voltage, in turn, will feed current into the next section, and $C_2$ will commence to charge up. This will feed current into $C_3$, and so the voltage will be transmitted along the feeder until we reach the far end. Here the feeder is terminated in an impedance, which we will assume to be a resistance (since we nearly always terminate a feeder in a tuned circuit which will have no reactance at resonance).

## Reflection.

What happens here depends upon the value of the resistance. Suppose it is infinitely high. The last condenser is unable to discharge through the load and will, therefore, force current *back along the feeder*, and energy will commence to surge back to the transmitting end. The wave will, in fact, be *reflected*.

Similar reflection occurs if we short-circuit the far end, for here the last condenser is not able to accept any charge, and the energy from the preceding section of the feeder has nowhere to go except back along the line, which it proceeds to do.

If the resistance is high but not infinite, it will accept some current from the last condenser, but not enough. The discharge of a condenser through a resistance takes time, and if the resistance will not discharge the condenser fast enough, the excess charge can only go back along the feeder. Thus, we have a partial absorption and partial reflection. If the resistance is too low the condenser discharges too fast, and again we have partial reflection.

There is clearly one particular value of resistance which accepts energy just as fast as it is being supplied by the feeder, and with this critical value *no reflection occurs.* **Energy is passed unhindered right along the feeder and is absorbed at the far end, with very little loss in transit.**

## Characteristic Impedance.

What is this value of resistance? It can be deduced by finding the rate at which the condensers charge and equating this to the rate at which the end section will discharge through the load. This involves Calculus, and the treatment is beyond the scope of the present work, but the result is surprisingly simple.

The conditions are fulfilled if the resistance is made equal to $\sqrt{(L/C)}$, where $L$ and $C$ are the inductance and the capacitance of the line,* or of a given length of line which comes to the same thing, since the total inductance and capacitance each depend upon the length. This is called the *characteristic* or *surge impedance*. It is the impedance of the line itself when the far end is open, and if we terminate it with a resistance of equal value, energy will be accepted from the line without reflection, irrespective of the length of the line.

The inductance of a pair of parallel wires (assuming the far end closed to complete the circuit) is

$$9{\cdot}21 \times 10^{-3} \log_{10} (d/r) \ \mu\text{H. per cm.},$$

where $d$ is the spacing of the wires and $r$ the radius.

* In a transmission line we have to consider two forms of loss

1. Loss of voltage due to resistance and inductance.
2. Loss of current due to leakage and capacitance.

The shunt current is dependent on the capacitance, and the larger this capacitance the greater the loss. The leakage current across the insulation may be expressed in terms of the *leakance*, which is the reciprocal of the insulation resistance. Hence, the larger the leakance the larger the current, which is a similar variation to that of the capacitance.

The complete expression, therefore, involves both $R$, the line resistance, and $G$, the leakance, and is of the form

$$Z = \sqrt{[(R + j\omega L)/(G + j\omega C)]}$$

At radio frequencies $L\omega$ is much greater than $R$ and $C\omega$ is much greater than $G$, so that the expression simplifies to $\sqrt{(L/C)}$.

Those readers who desire a more mathematical treatment should refer to *High Frequency Alternating Currents*, by McIlwain and Brainerd.

The capacitance between them is

$$\frac{10^{-4}}{828 \log_{10}(d/r)} \ \mu\text{F. per cm.}$$

Hence the characteristic impedance $\sqrt{(L/C)} = 276 \log_{10} (d/r)$ ohms. Over a wide range of practical values for $d/r$ this gives an impedance of the order of 500 to 700 ohms, and an average value of 600 ohms will be found to give results of the right order.

For short-wave work tubular feeders are often used, one wire running inside the other. The principle is the same, but the characteristic impedance is usually about 70 ohms.

## Matching the Line.

A feeder, therefore, provides a very flexible arrangement, which we can use for any frequency. All that is necessary is to ensure that the termination is correct, and we do this by matching the line to the output by high-frequency transformers, or by tapping the feeder across the tuned

FIG. 24. MATCHING THE INPUT AND OUTPUT BY AUTO-TRANSFORMERS

circuit in the usual way. The impedance of a parallel circuit at resonance is $L/CR$. This is usually very high, but by tapping down the coil we can make the effective primary impedance what we wish.

Fig. 24 shows a termination using a feeder of this type. Both the input and the output are tapped across the respective tuned circuits, so that the maximum efficiency is obtained. The input is matched to the feeder because, as is so often the case in communication work, the maximum efficiency is obtained when the internal and external impedances are equal Hence we obtain maximum energy

transfer into the feeder when it is suitably matched at the transmitting end.

The receiving end is matched to avoid reflection, which is another way of saying that the maximum energy transfer is obtained at this point as well.

It should be noted that the feeder only carries a current

FIG. 25. FEEDER HOUSE WITH AERIAL DOWNLEAD AT THE ATHLONE HIGH-POWER BROADCASTING STATION

The transmitting building is on the right

*(By courtesy of Marconi's W. T. Co.)*

necessary to transmit the requisite power, whereas the oscillating current in the circuits at either end is much greater than this, due to the usual resonant action. The process is the same as in the valve oscillator or amplifier, where the anode feed current is only large enough to maintain the losses, while the current oscillating in the tuned "flywheel" circuit is much greater.

Fig. 25 shows the aerial feeder at the Athlone station. The building in the foreground houses the aerial tuning inductance, which has to be large enough to handle several

hundred amperes, while the feeder from the main building only carries ten or twenty amperes and is thus quite light in construction.

In short-wave transmitters, it is often necessary to supply a number of aerials from the same feeder, and in this case special junction arrangements have to be made. The feeder is split into two, and each half into two again, and so on. At each point a matching transformer is arranged so that the impedance of the line viewed from either direction always appears to be the characteristic impedance $\sqrt{(L/C)}$.

## Standing Waves.

In a feeder having no losses the voltage at a distance $x$ from the transmitting end is shown on page 50 to be

$$E_x = E(\cos \beta x - j \sin \beta x)$$

where $\beta$ is the phase constant, which for a loss-less line is $\omega\sqrt{(LC)}$.

Since $\cos A + j \sin A = 1$, thism eans that the numerical value of the voltage is always the same, and equal to the voltage $E$, but the phase will be continually changing, making a complete rotation through $2\pi$ every time $\beta x$ becomes a multiple of $2\pi$. In other words there will be standing waves of *phase*, of wavelength $\lambda = 2\pi/\beta$.

This changing phase may or may not be of importance. In short-wave feeders where various aerials have to be supplied with current in the correct phase the length of the feeder has to be exactly chosen. Otherwise it is immaterial.

## Tuned Feeders.

Let us see what happens if the feeder is not correctly terminated. Consider a feeder open at the far end so that reflection occurs. The reflected wave will travel to the transmitting end, and again a reflection may occur if conditions are not exactly suitable. Consequently, instead of having just one wave travelling along the wire, we have a series of waves travelling backwards and forwards by successive reflections.

The currents produced by these various waves will all add up vectorially according to their relative phase, and

the result will be an uneven distribution, the current (and voltage) along the line being greater in some parts than in others. In other words we have a *standing wave*. The energy fed into the line will be entirely absorbed in building up an oscillatory current in the feeder, which will rise and fall just like the current in an ordinary tuned circuit. The system is, in fact, nothing but an elongated oscillating

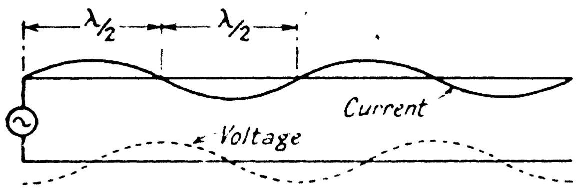



FIG. 26. STANDING WAVES ON A TUNED FEEDER

circuit, rather like the simple aerial discussed in the last chapter, but as there are two wires, the currents in which are equal and opposite at any given section, little or no radiation takes place.

Thus, we have alternate maxima and minima of both current and voltage. The points of zero voltage are called *nodes* and are separated by half a wavelength. This wavelength is that corresponding to the frequency of the current, i.e. $3 \times 10^8/f$ metres, and the current nodes occur at points of maximum voltage, midway between the voltage nodes.

Fig. 26 shows the distribution of voltage and current on an open circuited feeder which is an even multiple of half a wavelength long. Such a device is known as a *tuned feeder* or *Lecher wire*, and is sometimes used for measuring the wavelength of a transmitter. By finding the current "nodes" and measuring the distance between them, an accurate estimation of the wavelength can be arrived at.

The nodes are usually located by means of a shorting bar carrying a flashlamp or other indicating device which is moved along the feeder, maximum brilliance indicating a current node. The method is, of course, only suitable for short waves.

If the feeder is not an exact multiple of $\lambda/2$ the condition

of affairs is as if the generator in Fig. 26 were located at some intermediate point along the line. The voltage at the far end is still a maximum and the current at this point is clearly zero, as before. As we travel back towards the input end the voltage and current vary in the manner just described, giving nodes of either current or voltage every quarter wavelength until the generator is reached

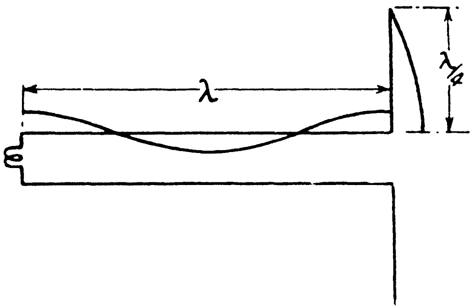



FIG. 27. TUNED FEEDER SUPPLYING A HALF-WAVE AERIAL

The voltage and current supplied by the generator are then determined by its position. If it were $\lambda/4$ to the right in Fig. 26, it would be supplying no voltage* and maximum current, and at any other point it would supply something less than the maximum. A feeder can thus provide a step-up action like an ordinary tuned circuit, the voltage at the far end being greater than the generator voltage, and this effect is sometimes utilized.

Tuned feeders may also be used to supply an aerial on short-wave systems, by merely extending the end of the feeder to form the aerial (see Fig. 27). The current and voltage relations in the system can now continue without any hindrance, and all that we have done is to increase the length of the aerial very considerably, but since the two wires of the feeder are running parallel the

* Assuming no resistance. Actually some voltage would be required to make up for losses.

feeder does not radiate as explained above, all the actual radiation coming from the aerial itself.

This form of tuned feeder, however, is only used where the run is short—not more than about one wavelength.

As distinct from the untuned transmission line, the resonant feeder has to carry the full oscillating current, so that the losses are greater and this restricts its usefulness to short runs.

The subject is further discussed in Chapter XII.

### Input Impedance of Tuned Feeder.

The input impedance of a tuned feeder clearly varies

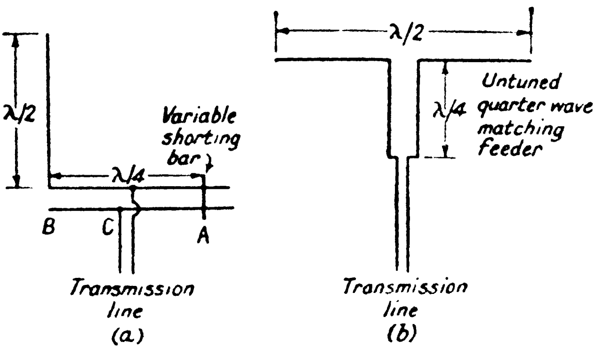



FIG. 28. MATCHING FEEDERS

between zero at a voltage node and infinity at a current node. The actual impedance (neglecting resistance) can be shown to be

$R_0 \cot \omega\sqrt{LC}\,x$ for a feeder with the far end open,

and $R_0 \tan \omega\sqrt{LC}\,x$ when the far end is short circuited,

$R_0$ being the characteristic impedance $= \sqrt{L/C}$.

### Matching Lines.

A feeder may thus be used to match two different impedances, the arrangement depending upon conditions. One example is shown at (*a*) in Fig. 28, where an untuned transmission line is being used to feed a half-wave aerial.

*The normal method of doing this is shown in Fig. 105, but* in the present method the aerial is voltage fed through a short length of tuned feeder made equal to an odd multiple of $\lambda/4$ and having its bottom end short circuited.

The impedance of the feeder thus rises from zero at $A$ to infinity at $B$, and the transmission line is tapped across at the point $C$ at which the impedance, as determined by the expression above (or, more likely in practice, by trial and error) is equal to the characteristic impedance of the transmission line.

Another method is to use a quarter-wave feeder in series with the aerial as shown at (*b*), Fig. 28. Under these conditions, it can be shown that no reflection occurs if the characteristic impedance of the matching line $R_m = \sqrt{R_1 R_2}$ where $R_1$ and $R_2$ are the impedances of the aerial and transmission line respectively.

The same device may be used to match any two impedances provided they are not too widely different. For proof, see page 61.

## Losses in Feeders.

We have neglected resistance throughout. In practice, resistance and other losses are inevitably present while radiation, though small, accounts for some loss. At radio-frequencies, however, the expressions deduced are not appreciably affected with normal lines.

The presence of losses will cause the current to be less than the theoretical value, but even this is only serious on long tuned feeders for the reasons already stated. In a large commercial station the matter receives some consideration and the feeders are designed to give the best efficiency.

With an open-wire feeder radiation loss decreases as the spacing of the wires is reduced, but resistance loss increases due to proximity effect, and a spacing which makes $R_0 = 600$ ohms is about the best. With concentric feeders radiation loss is negligible and the problem is one of conductor loss which is least if the radii of the external tube and central conductor have a ratio of 3·6 : 1. This is the

order of spacing used in the "coaxial" television cable and gives a value of $R_0$ equal to about 70 ohms.

## Transmission Equations.

We have discussed the properties of transmission lines in qualitative terms. A quantitative analysis involves mathematical treatment beyond the scope of this book, but it is desirable to state briefly the main expressions dealing with wave propagation.

By considering the line as composed of a number of small elements each receiving energy from its predecessor and passing it on to its successor, as explained at the beginning of this chapter, we can derive an expression for the voltage at any distance $x$ from the sending end. The complete expression is in two parts, one representing the forward wave and the other the reflected wave.

Assuming a correct termination, however, so that there is no reflection, the second term vanishes, so that for most practical cases we can write

$$v = V_s \varepsilon^{-\gamma x}$$

where $V_s$ is the voltage at the input or sending end and $\gamma$ is a factor known as the *propagation constant*.

We can evaluate $\gamma$ in terms of the line constants. The series impedance is $Z = R + j\omega L$. The shunt admittance is $Y = G + j\omega C$ (see footnote to page 42). It is then found that $\gamma = \sqrt{(ZY)}$. Note that since we are dealing with distributed constants $R$, $G$, $L$ and $C$ are per unit length.

We can rewrite the expressions in terms of resistance and reactive components.

$$\begin{aligned}
\gamma &= \sqrt{(ZY)} = \sqrt{[(R + j\omega L)(G + j\omega C)]} \\
&= \alpha + j\beta. \\
v &= V_s \varepsilon^{-\gamma x} = V_s \varepsilon^{-(\alpha + j\beta)x} \\
&= V_s \varepsilon^{-\alpha x} \varepsilon^{-j\beta x} \\
&= V_s \varepsilon^{-\alpha x} (\cos \beta x - j \sin \beta x).
\end{aligned}$$

Thus the wave is subject to a gradual decay, of exponential form, determined by the term $\varepsilon^{-\alpha x}$, accompanied by a progressive phase displacement determined by the second term involving $\beta$. Hence the factor $\alpha$ is called the ***attenuation constant***, while $\beta$ is known as the ***phase-shift constant***.

The evaluation of $\alpha$ and $\beta$ is cumbersome.* It is found that, provided $L\omega \gg R$,

$$\alpha = \tfrac{1}{2}[R\sqrt{(C/L)} + G\sqrt{(L/C)}].$$

The expression for $\beta$ is clumsy, but it may conveniently be written in terms of $\alpha$, in the form

$$\beta = \sqrt{[\alpha^2 - (RG - \omega^2 LC)]}.$$

Clearly $\alpha$ is independent of frequency but $\beta$ is not, so that the phase shift varies with the frequency. The actual phase shift at a distance $x$ is $\tan^{-1} \beta x$.

We have seen (p. 45) that the phase makes one complete cycle every time $\beta x$ becomes a multiple of $2\pi$. Hence the wavelength $\lambda = 2\pi/\beta$. The *velocity of propagation* along the line can thus be deduced from the fundamental relation $v = \lambda \times f$, so that $v = 2\pi f/\beta = \omega/\beta$.

If $R$ and $G$ are sufficiently small to be negligible by comparison with $\omega L$ and $\omega C$, so that the line may be considered as loss-free, $\alpha$ becomes zero and $\beta$ reduces simply to $\omega\sqrt{(LC)}$.

## Balanced Feeders.

If, in a two-wire feeder, one line is earthy, the other, being subject to the full variation of potential, may induce voltages in near-by wires and feeders giving rise to "cross-talk." To avoid this a balanced arrangement is often used such that as the potential on one feeder increases that on the other decreases by an equal amount. This can be arranged, for example, by earthing the mid-point of the transformer winding to which the feeder is connected or by some similar form of symmetrical connection.

## EXAMPLES III

(1) Calculate the surge impedance of a feeder 100 yd. long having two parallel wires $\frac{1}{8}$ in. diameter 6 in. apart.

(2) If the feeder of question (1) is to be connected to a tuned circuit of inductance 20 $\mu$H., capacitance 100 $\mu\mu$F., and resistance of 10 ohms, state the approximate tapping on the coil for correct matching.

* Because of the $\sqrt{}$ sign. The expression $\sqrt{(ZY)}$ can be written in the form $\sqrt{(p + jq)}$ quite easily. But $\alpha$ is *not* equal to $p$ nor $\beta$ to $q$ as the reader may quickly verify for himself.

# CHAPTER IV

## THE RADIO RECEIVER

### (a) Radio-frequency Amplification

The early forms of receiver consisted of a simple detector. Later, amplification was added after the detector, but as time progressed it became clear that for proper efficiency the detector should be operated under certain specific conditions. Methods were developed to amplify the high frequency signals before applying them to the detector, and the modern tendency is to arrange the greater part of the amplification prior to the detector stage.

**H.F. Amplifiers.**

The amplifier prior to the detector is thus of considerable importance and will be dealt with at length. We are concerned with voltage amplification only, in a radio receiver. The input voltage is very small—a few millivolts only—and our aim is to magnify these voltages as much as possible consistent with the limitations of adequate selectivity.

The fundamental principle of voltage amplification is that the external impedance shall be large compared with that of the valve, as explained in Volume I, Chapter XVII. The valve develops a voltage of $\mu e_g$ which forces current through the circuit comprising the internal resistance $r$ and the external load $Z$, as shown in Fig. 29 ($a$). The voltage developed externally is thus $e_2 = [Z/(Z + r)]\mu e_g$.

The denominator $Z + r$ is, of course, the vector sum but in r.f. amplifiers the anode load is usually in the form of a parallel tuned circuit, which is equivalent to a resistance $P = L/CR$, so that the gain of the stage (which is $e_2/e_g$) becomes simply $\mu P/(P + r)$.

If $P$ is large compared with $r$, as in the case of a triode, the gain tends to a limiting value $\mu$, but when a pentode valve is employed $r$ is usually large compared with $P$.

In such circumstances it is convenient to write the expression slightly differently, by multiplying both top and bottom by $r$. Then

$$v_2 = \frac{Zr}{Z+r} \cdot \frac{\mu}{r} \cdot e_g$$

The first part of this expression is the impedance of $Z$ and $r$ in parallel, while $\mu/r = g$, the mutual conductance of the valve. Hence the output voltage is that which would result from a current $ge_g$ flowing through $Z$ and $r$ in parallel.

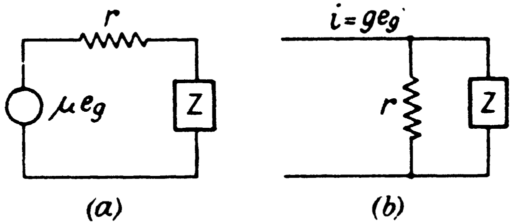



FIG. 29. EQUIVALENT VALVE CIRCUITS

The equivalent circuit may thus be drawn in the form shown in Fig. 29 (*b*) which is sometimes helpful. It shows, for one thing, that the valve resistance is in parallel with the load and thus adds to the losses, while if $r$ is much greater than $Z$, the stage gain becomes simply $gZ$.

## H.F. Transformers.

The use of a simple tuned circuit in the anode, however, is not necessarily the best. With a triode valve the total effective amplification may often be increased by tapping the anode across part of the coil (or using a transformer), while the damping due to the presence of the valve across the tuned circuit may also be reduced by using a transformer.

Consider the circuit of Fig. 30. The effect of the secondary on the primary may be calculated by the usual coupled circuit laws (see Chapter IX). Very approximately $Z_1 = M^2\omega^2/Z_2$, $Z_2$ being the secondary impedance, $M$ the mutual inductance between the windings and $\omega$ the angular frequency $= 2\pi f$, and since the secondary is tuned, $Z_2$ is

simply $R$, the secondary resistance. Hence $Z_1 = M^2\omega^2/R$. We have neglected the primary resistance and reactance, which is permissible in the present case because the "reflected" impedance of the secondary is many times greater.

The current in the primary, $i_1$, is thus

$$\frac{\mu e_g}{r + M^2\omega^2/R},$$

$r$ being the internal resistance of the valve.

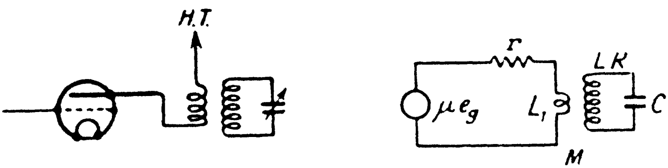



FIG. 30. SIMPLE HIGH-FREQUENCY TRANSFORMER AND EQUIVALENT CIRCUIT

The voltage induced in the secondary is $M\omega i_1$, and the voltage across the secondary coil is $(L\omega/R)e_2$

$$= \frac{L\omega}{R} \cdot M\omega \cdot \frac{\mu e_g}{r + M^2\omega^2/R}$$

$$= \frac{ML\omega^2\mu}{Rr + M^2\omega^2} \cdot e_g$$

The first part of this expression represents the total amplification of the stage, and by differentiating this it can be shown that the maximum gain results when $M^2\omega^2 = Rr$. This can be written $r = M^2\omega^2/R = Z_1$. In other words, the internal and external impedances are equal.

Consider a triode valve, having relatively low internal resistance, say 18 000 ohms. If the impedance of the tuned circuit is, say, 100 000 ohms, we shall extract practically the full gain out of the valve.

Let us now tap the anode half way down the coil (or use a 2 : 1 transformer). The equivalent impedance will be one quarter of the original amount (see next page), i.e. 25 000 ohms. This will only extract about 60 per cent of the available amplification from the valve, but we now have a 2 : 1

step-up giving an overall gain 20 per cent better than with the original arrangement.

A further reduction in primary turns would cause the amplification from the valve to fall off rapidly, and the step-up of the transformer will not be able to counteract this, so that the overall gain will be reduced. Hence there is an optimum value for the primary, given, as we have just seen, by the criterion $M^2\omega^2 = Rr$.

### Step-up Ratio.

The mutual inductance between two windings perfectly coupled without leakage is $\sqrt{(LL_1)}$, where $L$ and $L_1$ are the inductances. In practice, $M$ is less than this because the primary flux does not all link with the secondary, to allow for which we introduce a *coupling factor*, $k$, so that $M = k\sqrt{(LL_1)}$.

Clearly, if the coils are appreciably separated, $k$ will be small while, if they are closely coupled, $k$ will approach unity. In practice, with a primary wound over the low potential end of the secondary $k$ is about 0·8.

The criterion for maximum gain can thus be rewritten

$$\begin{aligned} Rr &= M^2\omega^2 = k^2LL_1\omega^2 \\ &= k^2L_1/C \text{ since } \omega^2 = 1/LC \end{aligned}$$

Hence $L_1 = CRr/k^2$

Now the step-up ratio $t$ is approximately $\sqrt{(L/L_1)}$ and the optimum value of this is obtained by substituting the value of $L_1$ just obtained. Hence

$$\text{Optimum step-up} = k\sqrt{(L/CRr)} = k\sqrt{(Z/r)}$$

where $Z$ is the "dynamic" impedance, $L/CR$, of the secondary circuit.

### Effective Primary Impedance.

We have seen that the equivalent primary impedance is approximately $M^2\omega^2/R$. We can write this also in terms of the step-up ratio, for

$$\begin{aligned} M^2\omega^2/R &= k^2LL_1\omega^2/R \\ &= k^2LL_1/LCR. \quad \text{But } L_1/L = 1/t^2 \\ \therefore Z_1 &= k^2 \,.\, L/CRt^2 = k^2 \,.\, Z/t^2. \end{aligned}$$

Thus the effective primary impedance is inversely proportional to the square of the step-up ratio, as was mentioned previously.

## Screen-grid Valves.

With a triode the optimum value of $t$ is usually between 2 and 3. Modern equipment, however, almost invariably uses screened valves in which $r$ is anything from 0·25 megohm upwards. With such conditions $t$ becomes fractional.

For example, if $r = 500\,000$ ohms and the circuit has a dynamic impedance of 100 000 ohms (which is about the average for commercial receivers) the optimum step-up, with $k = 0{\cdot}8$, becomes 0·36, representing a step-*down* of 2·8 to 1.

Such a transformer, however, is rarely used from considerations of selectivity, because of the effect of the valve on the tuning of the circuit, and in fact the circuits are often used with a step-up ratio, which is still further from the optimum as regards gain, but gives improved selectivity.

Under these conditions, with an external impedance much less than that of the valve, the expression for the valve amplification reduces to $(Z/r)\mu$. Since $\mu/r = g$, the mutual conductance, this equals simply $gZ$. The total gain is, of course, $t$ times as great, where $t$ is the step-up ratio, if any.

Let us examine the behaviour of an h.f. transformer under these conditions. The gain from the valve will be $gZ_1 = gk^2Z/t^2$ and the overall stage gain will be $gk^2Z/t$.

The additional loss due to the valve (the valve damping) can be calculated from the same coupled circuit laws, used this time in the reverse direction. Thus the additional secondary resistance due to the presence of the valve load $r$ across the primary is

$$M^2\omega^2/r = k^2LL_1/LCr = k^2L/Crt^2 = k^2LR/CRrt^2$$
$$= (k^2/t^2)\,(Z/r)\,R.$$

If $k = t = 1$ and $Z = r$, this reduces to $R$. In other words with a tuned anode circuit having a dynamic impedance equal to the valve resistance the effective resistance

of the tuned circuit is doubled by the valve damping, which we know to be correct.

The formula also shows that the valve damping is directly proportional to $Z/r$ so that the better the circuit the more severe will be the effect of the valve on the tuning, while it is inversely proportional to $t^2$, giving a rapid improvement in conditions as the step-up ratio increases, while the gain only falls off directly with $t$.

## H.F. Pentodes.

A screen-grid valve can be made to possess either a moderate a.c. resistance or a very high one—over a megohm being common practice—with practically *the same mutual conductance*.

Any increase in $r$ reduces the damping imposed on the tuned circuit in direct proportion, while the gain remains at $gk^2Z/t$ since $g$ is unaltered. In fact, it no longer becomes necessary to tap down the coil and a plain tuned anode circuit may be used giving the full gain $gZ$ with adequate selectivity.

The type of valve used in the modern receiver, therefore, is of this high impedance type, and it has become so efficient that large anode swings are practicable. In such circumstances the secondary emission kink in the conventional tetrode characteristic seriously restricts the available swing and pentodes are customarily employed. These valves contain a suppressor grid (at cathode potential) between screen and anode to suppress the secondary emission as explained in Vol. I, Chapter XII.

Use is also made of "critical distance" tetrodes, though these are more employed in audio-frequency stages as explained on page 124.

## Selectivity.

The selectivity of a circuit is the measure of the frequency discrimination exercised by the tuning. It may be expressed in various ways, one being the ratio of the voltage developed across the circuit to that at resonance. This is

$$S = \frac{\text{voltage at frequency } f}{\text{voltage at resonance}}, \text{ which can be shown to be}$$

$$= 1/\left[1 + \Delta + jQ\Delta\left(\frac{2+\Delta}{1+\Delta}\right)\right]$$

where $\Delta = (f - f_r)/f_r$

$Q = 2\pi f_r L/R$

If $\Delta$ is small, this can be simplified to

$$S = 1/(1 + 2jQ\Delta) = 1/\sqrt{1 + 4Q^2\Delta^2}$$

In this expression, the value of $R$ used in calculating $Q$ must be the total circuit resistance inclusive of the damping produced by any shunt impedance. If the circuit is shunted by a resistance $P$, the effective series resistance becomes $R = R_0 + L/CP$, $R_0$ being the original resistance. (See page 102.)

In practice there are always two shunt loads on an intervalve transformer. One is the input resistance of the following valve which includes any grid leak present. At long and medium wavelengths the grid resistance is due mainly to Miller effect, as discussed on page 98.

The second, and usually more important effect is that of the preceding valve, and, as we have seen, the conditions for maximum gain occur when the effective primary impedance is equal to the valve resistance, in which circumstance the effective $Q$ of the circuit is halved.

As just explained, however, the modern r.f. pentode has an impedance many times higher than that of the external circuit so that valve damping is of minor importance. The circuit loss is, in fact, the major factor and this cannot be reduced too much or the stage gain becomes inconveniently high, causing instability. Increased selectivity is obtained by the use of coupled circuits as explained later.

## Variation of Gain with Frequency.

The amplification will be seen to depend on the frequency, and will therefore vary as the circuit is tuned over the frequency range. This is an inherent defect in "straight" receivers. The effect may be offset by using special primary circuits which transfer more energy to the secondary at the lower frequencies.

Such a circuit is shown in Fig. 31. The energy from the anode circuit is introduced into the secondary partly through the mutual inductance $M$, and partly directly by virtue of the voltage developed across $C_1$ by the primary current flowing through it. If the frequency falls the inductive voltage falls but the capacitive voltage increases and by correct proportioning a nearly constant voltage transfer can be obtained. /

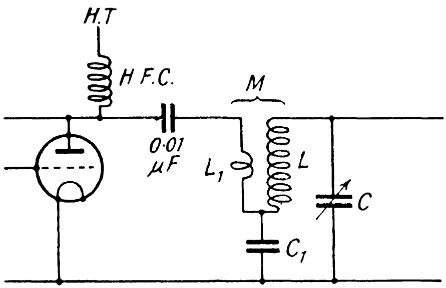



FIG. 31. "CONSTANT-COUPLED" CIRCUIT

The voltage across the whole secondary circuit is greater than the induced voltage because of the magnification of the circuit ($L\omega/R$), but this also falls off as the wavelength increases. The coupling can be arranged to compensate for this as well, giving constant voltage across $C$.

The capacitive energy transfer may be achieved in other ways, e.g. by a small top capacitance coupling, while similar measures may be adopted in the aerial coupling circuits as explained on page 143.

## Stability.

It is important in any amplifying stage to avoid any unauthorized feed-back of energy from the anode to the grid circuit. If this feed-back is sufficient in magnitude and in the right direction, continuous oscillation will result, as in the case of a simple valve oscillator. If the feed-back is in the reverse direction the gain of the stage will be appreciably reduced.

There are two principal methods by which feed-back takes place. One is due to direct magnetic or electrostatic coupling between portions of the circuit or the wiring. Currents are induced from one circuit to the other, and, owing to the very high amplification which can be obtained (a gain of 100–150 is quite normal in a modern receiver), this feed-back has only to be very slight in order to produce a marked effect.

The second cause of instability is the internal capacitance between anode and grid of the valve itself. The magnitude

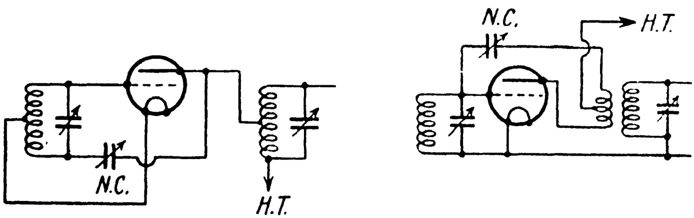



FIG. 32. TWO FORMS OF NEUTRALIZED CIRCUIT

and direction of this internal feed-back depends upon the nature of the anode impedance, as explained on page 98. When the anode circuit is nearly in tune the feed-back is positive and may easily be sufficient to produce continuous oscillation. In a triode valve this is a very serious difficulty, and it is overcome by the use of neutralizing circuits in which energy is fed from the output to the input, through a special circuit, in opposition to that which passes through the valve. If these circuits are symmetrically designed, the adjustment remains adequate over a wide range of frequency, such as would be covered by the normal tuning operation.

Fig. 32 shows two satisfactory forms of circuit. The first circuit, due to Rice, is very symmetrical, but suffers from the disadvantage that only half the full voltage is supplied to the grid of the valve. In the second, due to Hazeltine, a separate neutralizing winding is used, exactly similar to the primary winding, but connected in opposite phase. It is essential that the primary and neutralizing winding shall be very tightly coupled, and the usual practice is to wind one

over the other. Modifications of these circuits are used, but they all employ one or other of these two basic forms.

## Shielding.

The alternative to this is to reduce the capacitance in the valve itself to such small limits that it is not troublesome under normal conditions. The customary r.f. pentode contains an internal screen for this purpose, as explained in Vol. I, and is almost invariably employed for this reason. But, having removed this source of feed-back, it becomes all the more important to eliminate any other coupling between the circuits. Considerable feed-back can arise from interaction between the tuning coils, to avoid which it is customary to enclose them in metal shields or cans. These shielding boxes are made of copper or aluminium, and operate as follows—

The magnetic fields generated by the coils induce eddy-currents in the material of the shielding. These eddy-currents in turn produce magnetic fields of their own which, by Lenz's law, are in opposition to the original field. Consequently, the field outside the can is negligibly small. The process obviously involves a loss of energy, since the eddy-currents circulating in the material of the can must absorb power. Also, the magnetic field set up by the shield reduces the effective magnetic field within the coil, so that a coil has a lower inductance inside the can than out of it. By suitable design the loss of efficiency can be made relatively small.

It is important that each circuit shall have its own screen. One common screen for all the circuits, even if it is partitioned off, is unsuitable, as circulating currents in the screen may induce voltages from one circuit to another and so defeat the whole object of the screening.

Electrostatic coupling, due to the difference of potential between the various parts of the circuit, is minimized by interposing a simple metal shield between the requisite portions. The material in these shields is of good conductivity (usually copper or aluminium, *not* iron) in order to minimize the loss produced by the eddy-currents.

The general methods adopted in modern practice are illustrated in Fig. 33. Note particularly the condenser, in which three separate sections are mounted on a common spindle. The individual sections are made as nearly identical as possible and, in some cases, the end plates are slotted radially at intervals, allowing them to be bent slightly

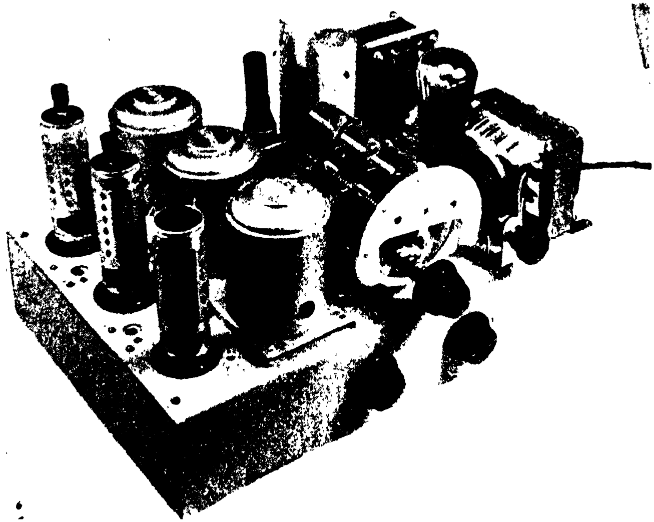


FIG. 33. TYPICAL MODERN CHASSIS OF A "STRAIGHT" RECEIVER WITH TWO HIGH-FREQUENCY STAGES

Note the screened coils and the "gang" condenser, comprising three condensers on a common spindle

in order that the capacitances at fixed points in the scale may be made absolutely identical. The tuning coils are similarly matched as regards inductance and final adjustment is made on "trimmers," which are small condensers mounted in parallel with each condenser section to allow for small differences in the stray circuit capacitances.

For the reason already given, any electrostatic screening, including the chassis, *must not be allowed to carry any h.f. currents.* Otherwise coupling will be introduced between the circuits. The chassis should not be used as an earth line.

A separate heavy gauge bus-bar should be run between the appropriate points, this bus-bar being connected at one point only to the chassis, preferably at the actual earth terminal.

## Grid Input Impedance.

Mention has already been made of the fact that the grid circuit of the valve following a tuned circuit does not present an infinite impedance. Apart from any grid leaks and stray capacitances, the valve itself introduces loss from two causes. These are—

(*a*) the *Miller* effect, due to internal feed-back;

(*b*) the *Ferris* effect, due to the transit time of the electrons.

The former effect is a result of the feed-back just discussed through the internal capacitance between anode and grid. The subject is discussed in detail on page 98, where it is shown that the grid-cathode capacitance is increased by $C_{ga}\,(1 + A \cos \theta)$ and an additional resistance is introduced across grid and cathode $= - 1/\omega C_{ga} A \sin \theta$.

$C_{ga}$ is the grid-anode capacitance, $A$ is the amplification of the valve and $\theta$ the phase angle.

At resonance, $\theta = 0$ and $\cos \theta = 1$. Under any other condition $\cos \theta$ is less than 1. Hence as the anode circuit is tuned the input capacitance rises to a maximum and falls away again. This effect distorts the tuning of the input circuit and is known as "pulling." It is only serious with triode amplifiers.

The shunt resistance varies more. If the anode tuning capacitance is too large, $\sin \theta$ is negative and the grid resistance is positive, i.e. the input circuit is subjected to additional damping. At resonance, $\sin \theta = 0$ and the input resistance is infinite, so that there is no effect on the circuit other than the additional capacitance referred to above. If the anode circuit is inductive the input resistance becomes negative, causing a reduction in damping which, if sufficient, will cause continuous self-oscillation.

This is the reason for adopting neutralizing, or using screen-grid valves. The effect is only serious with triodes as the reader may verify for himself, but it still is present

even with h.f. pentodes because although $C_{ga}$ is of the order of 0·001 $\mu\mu$F. the stage gain is so much higher.

The two effects combine to produce a distortion of the tuning curve. At frequencies above resonance the additional damping reduces the response, while below resonance the response is maintained beyond the true resonance point by the negative input resistance. The result is a flat-topped resonance curve slightly displaced from the true resonance position. With screen-grid valves the displacement is small.

The second effect, due to transit time, was described in detail by W. R. Ferris, *Proc. I.R.E.*, Jan., 1936. When the time of oscillation becomes comparable with that taken by the electrons to move between the electrodes of the valve, a serious loss occurs. The effect is only evident at frequencies above about 10 megacycles/sec. ($\lambda$ = 30 metres) but is very troublesome at ultra-short waves. Special "acorn" valves having very small clearances have been introduced for such frequencies. The subject is discussed in detail in *Short Wave Radio* (Pitman) by the author.

## Multi-stage Amplifiers.

For simple receivers a single h.f. stage is sufficient with modern valves, but more powerful receivers use two or three stages. When this is done, further precautions have to be taken to maintain stability. The high order of amplification obtainable renders it imperative to avoid any coupling between output and input.

Attention to layout will usually result in a successful amplifier, but it is sometimes necessary to shield "hot" leads by enclosing them in a metal tube or similar covering which is connected to earth.

A second form of trouble with multi-stage amplifiers arises from common impedances. In Fig. 34, for example, the anode current of both valves has to run through common impedance $Z$ (which may be accidental in the circuit). The voltage developed across this impedance by the anode current of the second valve is thus automatically introduced into the anode circuit of the first valve, where it will produce feed-back, causing either an increase or decrease in amplification according to the phase. This difficulty is overcome

by decoupling the various leads. A filter comprising an h.f. choke or resistance with a condenser bypass is inserted in each lead, as shown in Fig. 35. The h.f. currents are

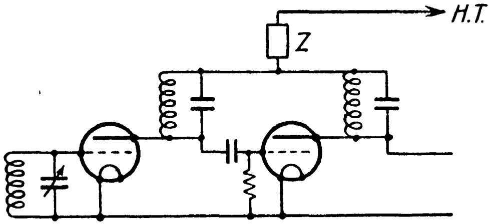



FIG. 34. THE COMMON IMPEDANCE Z IN THE HIGH-TENSION LEAD WILL CAUSE COUPLING BETWEEN THE VALVES

thus bypassed to earth without going through the battery or power-supply unit.

## Residual Signal.

The layout of an amplifier is also of vital importance from the point of view of selectivity. From the knowledge of the constants of the circuits it is possible to calculate the resonance curves (allowing for valve damping and similar factors) and arrive at an estimate of what is called the *adjacent channel selectivity*, which means the reduction in strength likely to be obtained on a signal relatively close in frequency to (only a few kc/s. away from) the wanted station.

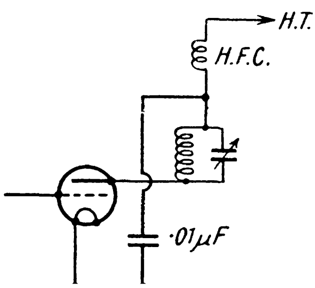



FIG. 35. HIGH-FREQUENCY DECOUPLING CIRCUIT

A still more important requirement, however, is that the selecting action shall continue to be progressive and even with good circuits this may not necessarily be the case. Once the relatively large resonant gain has been lost,

i.e. at frequencies 20 kc/s. or more off resonance, the circuit gives a small but nearly equal response to any frequency within quite a wide range.

With a powerful local station close at hand, the residual voltage even well away from the tuning point may be quite large and it is necessary therefore to ensure that each circuit continues to exercise its full selecting action.

This is partly a matter of circuit design, for the use of a high $L/C$ ratio ensures a fairly steep "skirt" to the resonance curve, but is equally a matter of layout. Suppose we have two circuits, separated by one or more amplifying valves, each containing a short length of common lead, perhaps in the earthy or low potential side where it might be assumed unimportant, then two effects will occur.

The currents in the last stage will set up voltages due to the impedance of this common lead, which will be reintroduced via the first circuit into an earlier portion of the amplifier. This will produce reaction which may either increase the gain (and in the limit produce instability) or reduce the gain and cause the set to be less sensitive than it should be.

The second and even more pernicious effect is that residual signals from a powerful station, partly filtered by the first stage, have a direct entry into the later stage without having to go through the filtering action of the intervening stages, so that the *remote channel selectivity* suffers seriously.

Such matters require careful attention if an amplifier is to behave properly. A good test is to see whether the amplifier "cascades" as it should. If the gain of two adjacent stages is $n$ and $n_1$ the gain of the two together should be $nn_1$. If it is not, then some common coupling is present and it should be located and removed.

## Noise Level.

In broadcast technique the number of stages is limited, mainly from considerations of cost. Where this does not apply, it is possible to increase the number of stages and to obtain thereby an improved performance principally in the direction of increased selectivity. It is necessary to

limit the amplification per stage because there is a maximum overall amplification beyond which useful results cannot be obtained.

This limit is set by the noise level in the amplifier which manifests itself as an indeterminate background which is amplified at the same time as the signal, and it is clearly necessary that the signal-to-noise ratio shall be such as to enable the signal to be clearly distinguished. The cause of background noise is threefold.

(1) *Atmospheric Disturbances or Interference.* Atmospherics are random waves which emanate from natural sources. They are in the form of damped waves of so short a duration as to be practically instantaneous, and they appear indiscriminately over the whole frequency spectrum. Their effect is discussed in Volume I, Chapter XXVIII, and the only satisfactory method of combating them is to use a combination of sharp selectivity and directional methods.* The latter is the more effective, and if the transmission can be made strongly directional so that a large signal is obtained at the receiver in the first place, and the receiver is also made directional so that it only receives within a narrow angle containing the wanted signal, then the atmospheric disturbance can be reduced many times.

Interference from electrical plant is a local disturbance and is more troublesome in broadcasting than in commercial practice. The commercial receiving station is located away from such sources and adequate suppression is fitted to any machinery which has to be in the vicinity.

With broadcast technique this is not always possible, and additional devices have to be used at the receiver to minimize the disturbance. The subject is discussed further in Chapter VIII.

(2) *Shot Noise.* This is a noise arising from uneven emission of the electrons in a valve. If adequate emission is provided from the cathode, however, this is continually surrounded by a space charge which acts as a reservoir so that the effect is reduced considerably, and in fact at medium and long waves it is negligible in comparison with the thermal agitation discussed in the next section. At

* Or to employ frequency modulation.

short waves and with frequency changers which have low conversion conductances, however, shot noise is predominant. A convenient treatment is that due to Dr. James of the G.E.C. Research Laboratories.

The equivalent grid noise due to shot effect can be shown to be given by

$$v^2_s = 2keF\,\Delta f\, i_a/g^2$$

where $e$ is the charge on an electron.

$i_a$ is the anode current.

$g$ is the mutual conductance.

$\Delta f$ is the frequency band of the receiver.

$k$ is a constant, and

$F$ is a factor dependent on the operating conditions.

With a temperature-limited cathode $F = 1$, while with a valve such as a triode the space charge, acting as a reservoir as already described, considerably reduces the noise, and $F$ has a value of about 0·05.

With a screen-grid valve the presence of the screen overcomes the space charge to a large extent, while the random distribution of electrons between anode and screen further increases the noise. In other words, although the steady anode and screen currents bear a fixed relation for any given operating conditions, the instantaneous values are fluctuating slightly, and as a result of these various factors $k$ increases to a figure of the order of 0·25 to 0·3.

A screened valve, therefore, produces five or six times the noise of a triode, while with a frequency changer the conditions are still worse, as explained on page 84.

The criterion of goodness, from a noise point of view, is determined, for a given class of valve, by the ratio $g^2/i_a$. The higher this can be made the less will be noise.

(3) *Thermal Agitation*, sometimes called Johnson noise. This arises from the movement of the electrons in the material of the conductors. Both this and the shot noise are only important in the first stage of the amplifier where

they are followed by the full amplification of the remaining stages. In the case of thermal agitation, the impedance across the grid and cathode of the first valve develops a noise which is proportional to the effective resistance and to the absolute temperature. As is to be expected, the noise is random and is distributed over the whole frequency spectrum, so that a receiver with a wide band width inherently possesses more background noise than a sharply selective receiver, which is a further argument for restriction of the band width to the minimum necessary to fulfil the particular requirements.

The magnitude of the voltage produced can be calculated from the formula—

$$E^2 = 4kT\int_{f_1}^{f_2} R\,df$$

where $k$ = Boltzmann's constant = $1{\cdot}374 \times 10^{-23}$

$T$ = absolute temperature = 273 + temp. in °C.

$R$ = resistance component of input impedance

and $f_1$ and $f_2$ are the frequency limits of the receiver.

This is a general formula which assumes that the resistive component of input impedance is not constant but varies with frequency. In the more usual case the resistive component is constant, in which case the formula reduces simply to—

$$E^2 = 4kTR\,(f_2 - f_1)$$

It should be noted that a tuned circuit behaves as a high resistance = $L/CR$ and that this dynamic resistance may be substituted in the formula. The magnitude of the effect may be easily assessed. For example, a tuned circuit having an effective resistance of 0·25 megohm at a temperature of 300° K., followed by an amplifier with a 10-kilocycle band width would develop 6·4 microvolts. To provide a satisfactory signal-to-noise ratio, therefore, we must have a signal of 50 or 60 microvolts which accounts for the statement made on page 39, that the smallest commercial signal is of the order of 10 microvolts per metre.

Moreover, the ordinary detector operates at a voltage

of a few volts, so that the maximum r.f. gain which can be provided to the detector is less than 100 000, a surprisingly low figure and one which can easily be exceeded with modern technique.

For further information on these points the reader is referred to a paper by E. B. Moullin and H. D. M. Ellis on "The Spontaneous Background Noise in Amplifiers due

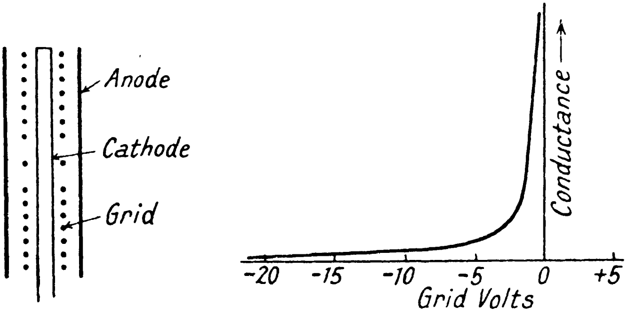



FIG. 36. ILLUSTRATING CUT-AWAY PORTION IN CENTRE OF GRID TO GIVE VARI-MU ACTION AND THE TYPE OF CHARACTERISTIC OBTAINED

to Thermal Agitation and Shot Effects," *Journal I.E.E.*, 1934, Vol. 74, page 323.

## Gain Control—Vari-mu Valves.

The signal at the output end of the receiver, either in the form of audio-frequency output to a loud speaker or telegraphic signals to a recorder, must obviously be capable of control. The necessity for keeping the voltage applied to the detector within certain limits has already been mentioned, and this leads to the need for some means of controlling the amplification of the h.f. stages.

A most satisfactory way of achieving this is by using what are known as *vari-mu* valves. These are provided with a specially-designed grid so that the slope decreases progressively as the grid bias is increased. Consequently, the effective amplification is under easy and simple control by merely altering the grid bias applied to the system.

The valves are usually tetrodes or p.f. pentodes and are used just as an ordinary valve. Frequency-changing valves are also usually provided with this form of control.

The subject is discussed fully in a paper by Ballantine and Snow entitled "Reduction of Cross Talk in Radio Receivers," *Proc.I.R.E.*, Vol. 18, p. 2102, December, 1930.

Theoretically, the valve should obey an exponential law so that the rate of change of slope at any point is proportional to the slope itself at that point. This results in a valve having a rather low maximum conductance and, therefore, a compromise is adopted, by omitting some of the wires in the centre of the grid, as shown in Fig. 36. Then, when the major part of the electron emission is cut off by the negative voltage on the grid, there is still a "hole" through which some electrons can flow, and since the grid at this point is a wide-mesh one, the effective amplification factor is quite small. By suitable design a smooth variation of slope can be obtained.

## Cross Modulation.

Any ordinary valve will, of course, give decreased amplification as the grid bias is increased, but the variation is not uniform, and this gives rise to a pernicious effect known as *cross modulation*. Let us assume that we have an interfering signal several times as strong as the wanted signal. Normally we should pass the signals through still further tuned stages at each of which the ratio of wanted to unwanted signal would improve, so that in the end the inference would be eliminated (see page 91).

In the early stages, however, the strength of the interfering signal may well be large enough to swing the valve over so extended a portion of its characteristic as to introduce serious distortion. Under these conditions the effective gain of the valve will depend on the strength of the interfering signal so that the output of the wanted signal will be controlled by the modulation of the interfering signal in addition to its own legitimate modulation.

This cross modulation, once introduced, *cannot be removed by subsequent tuning*. It must therefore be avoided at the start, firstly by the best practicable pre-selection in the

tuned circuits prior to the first valve, and secondly by ensuring that the said valve will handle a really large input without serious distortion. A vari-mu valve is so designed, particularly with the bias run back. This, of course, reduces the gain which may result in an increased relative noise level so that, as always, the designer has to adopt the best compromise to suit his particular requirements.

Other methods of radio-frequency gain control are sometimes used, such as variation of screen voltage, alteration of aerial coupling, etc. These, however, are of limited application and the majority of modern circuits use the vari-mu technique.

## EXAMPLES IV

(1) Calculate (*a*) the mutual inductance between primary and secondary, and (*b*) the approximate step-up ratio, for optimum gain in a h.f. transformer having a 165 $\mu$H. secondary following a valve of 18 000 ohms resistance when

(i) $f =$ 1 200 kc/s. $R =$ 13 ohms.
(ii) $f =$ 600 kc/s. $R =$ 5 ohms.

(2) With a transformer designed to suit (ii) above, calculate the additional resistance in the secondary due to the valve at the two frequencies.

(3) Calculate the approximate gain from a screen-grid valve having an amplification factor of 1 000 and an internal resistance of 300 000 ohms, with the transformer designed to suit (ii) above, at the two frequencies.

## CHAPTER V

# THE RADIO RECEIVER

## (b) Superheterodyne Receivers

The superheterodyne receiver is used very considerably to-day. As explained in Volume I, the procedure here is to convert the incoming signal to a different frequency by mixing it with a suitable local oscillation. A modulated wave is then produced which varies at a frequency corresponding to the difference between the incoming oscillation and the local oscillation. For this difference frequency to be detected, it is necessary that the combined current shall be rectified or passed through some form of non-linear device. The difference tone is then accepted by the use of tuned circuits, and is suitably amplified until the voltage is sufficient for the particular detector in use.

The advantage of this procedure is that the tuned circuits employed are all fixed-tuned (except the aerial and oscillator). Hence, high-efficiency circuits can be used giving high gain and improved selectivity without the troubles attendant on the use of ganged h.f. stages.

In particular, the amplification in a "straight" set varies with the setting of the tuning condenser, since the dynamic impedance $L/CR$ becomes less as $C$ is increased. This difficulty is obviated in a superhet, which gives uniform gain *and selectivity*, an important advantage.

The difference tone is still a radio frequency, the actual value depending largely upon the radio-frequency signals to be received in the first case. For instance, for broadcast operation where the frequency ranges from 1 500 to 150 kc/s., the intermediate frequency may be chosen at about 110 kc/s. On the other hand, for short-wave reception where the frequencies to be received are of the order of several megacycles, an intermediate frequency of 400 to 500 kc/s. is often employed, while lower or higher frequencies are used for special purposes. In any case, the intermediate

frequency is much higher than the relatively low speech modulations, which are therefore transferred through the amplifier to the detector, unaffected by the change in the carrier frequency.

## I.F. Amplifiers.

The construction of intermediate-frequency amplifiers is similar to ordinary radio-frequency practice, except that variable tuning condensers of the ordinary pattern are not required. Instead, small pre-set condensers are used usually mounted inside the shielding boxes which contain the coils. Also, in order to obtain selective tuning without loss of upper frequencies, band-pass circuits are used, or, alternatively, the individual circuits are slightly mistuned.

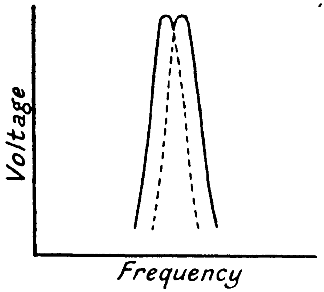



FIG. 37. OBTAINING FLAT TOP BY MISTUNING

The usual practice is to tune both anode and grid circuits and to couple these two together magnetically. The coupling is adjusted by altering the spacing between the coils, and this may be chosen to give the critical coupling which just causes the combined resonance curve to double hump, or else the coupling is set slightly weaker than this, but the two circuits are deliberately mistuned so that they give an effective broad top to the resonant curve as shown in Fig. 37. The circuit is a particular form of band-pass filter and is analysed further in Chapter IX.

Both the tuning circuits are usually mounted in the same can with small holes through the top by which the individual tune can be altered, and in many cases the actual resonance curve is traced by a cathode-ray oscillograph and the adjustment made until it obeys a pre-determined shape. Fig. 38 shows a typical intermediate-frequency transformer.

The stage gain is usually more than can be conveniently obtained with tuned radio-frequency stages for two reasons.

(*a*) The resistance of a coil of given inductance decreases with frequency so that although the relatively low inter-

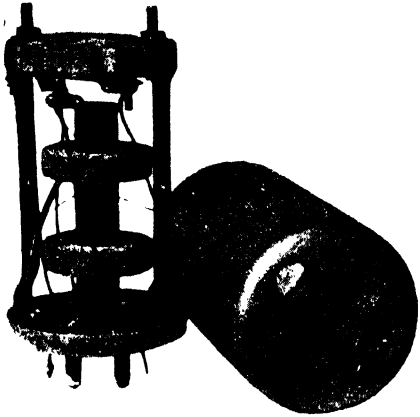


FIG. 38. TYPICAL INTERMEDIATE-FREQUENCY TRANSFORMER WITH COVER REMOVED

The tuning condensers are of the compression type, and are mounted on the steatite top portion. They are tuned with a screwdriver inserted through the holes seen in the cover

mediate frequency necessitates inductances of 10 000 $\mu$H. or more, the ratio $L/R$ is far higher than with a normal signal-frequency coil.

(*b*) The fixed tune enables the circuit to develop its maximum gain. With a signal frequency stage the set must be designed to accommodate the maximum gain, which occurs towards the bottom of the tuning scale. Over the majority of the scale, therefore, the gain is appreciably less than the maximum.

The fixed tune also enables the best ratio of $L/C$ to be adopted. Theoretically the gain increases as this ratio is made larger, but this is offset to some extent by the increase in resistance.

## I.F. Stage Gain.

Generally speaking, the dynamic resistance of an i.f. transformer circuit is between 0·25 and 0·5 megohm, giving gains of several hundred. With such circuits the full

formulae must be used, for $Z$ is no longer small compared with $r$.

The overall gain with a critically adjusted band-pass transformer is roughly half the gain from the valve itself. As explained in Chapter IX, critical coupling occurs when $M^2\omega^2 = R_1R_2$, $R_1$ and $R_2$ being the resistances of primary

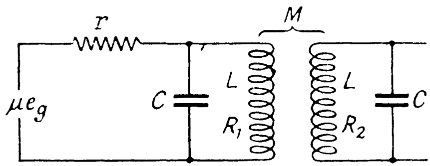



FIG. 39. TUNED I.F. TRANSFORMER CIRCUIT

and secondary circuits, usually equal. Under these conditions the performance of the circuit can be calculated fairly easily.

Consider the circuit of Fig. 39. We have assumed the primary and secondary to be similar. The dynamic impedance of the primary $Z_1$ will be $L/CR_1$ and this by itself will extract from the valve a gain of $\mu Z_1/(r + Z_1)$.

To find the voltage across the secondary we can work in two ways. Either we can evaluate the current in the primary, determine the voltage induced in the secondary and from this deduce the voltage across the whole secondary circuit, or we can replace the two tuned circuits by an equivalent single circuit.

The method of arriving at this is explained in Chapter IX. The effective primary resistance becomes

$$R_1' = R_1 + M^2\omega^2/R_2.$$

We have, however, assumed critical coupling so that $M^2\omega^2 = R_1R_2$. Hence $M^2\omega^2/R_2 = R_1$ and $R_1' = 2R_1$. In other words the presence of the secondary has doubled the effective resistance of the primary.

There is normally a reduction of primary reactance due to the secondary but, since the secondary is tuned, $X_2 = O$, and this effect does not enter into the calculation.

The effective dynamic impedance is thus $L/2CR_1 = Z_1/2$, and the gain becomes as $\mu Z_1/2(r + 2Z_1)$. If $Z_1$ were small

compared with $r$ this would be exactly half the previous value. Actually it is a little less.

It should be noted that $R_2$ must take into account the effect of any shunt resistance across the secondary. If the following valve is an amplifier there will be little loss on this account, but if the transformer is feeding a diode the damping will be quite heavy. The point is discussed further in the next chapter. It is sometimes sufficiently great to warrant tapping the detector some way down the coil.

The effective selectivity of the whole circuit may be arrived at by similar treatment. In this case the equivalent parallel circuit of Fig. 29 (*b*) is more convenient. We reduce the whole circuit to a single equivalent circuit allowing for

(*a*) The reflected secondary losses referred into the primary circuit.

(*b*) The damping due to the valve resistance $r$ in parallel with the circuit.

The selectivity may then be assessed as shown on page 57.

## Frequency Changers.

The conversion of the incoming radio frequency to the intermediate frequency is a matter of some importance, and numerous methods have been employed from time to time. Where considerations of space and cost permit, the most satisfactory method is to use two separate valves for the purpose, one to generate the oscillation and the other to do the mixing. The oscillator valve is a simple arrangement using one or other of the well-known oscillating circuits. It is desirable to include a grid condenser and leak in order to maintain the oscillation reasonably constant, and also to avoid excessive grid swings which generate harmonics in the output.

Harmonics are to be avoided because they will heterodyne with stations other than the wanted one and introduce audible whistles, particularly if they happen to coincide with the frequency or the harmonic of a powerful local station. It is desirable, therefore, to use a good oscillator circuit having a high $Q$* and a weak reaction coupling.

* The symbol $Q$ is used to denote the magnification of a coil $= L\omega/R$.

The self-limiting action of the grid condenser and leak is only partially effective. The circuit will tend to oscillate more readily at low values of tuning condenser and the reaction may be so great that "squegging" occurs. This is

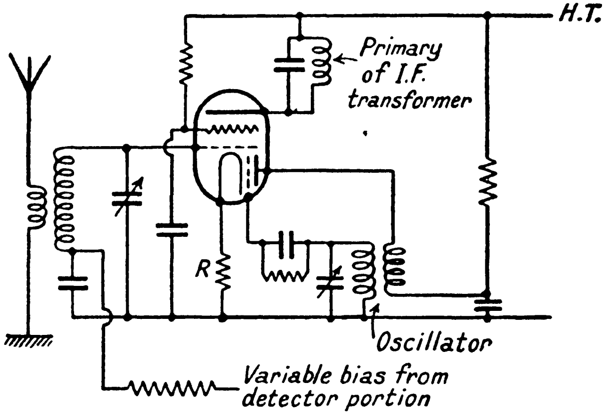



FIG. 40. TRIODE-PENTODE FREQUENCY-CHANGER CIRCUIT

The suppressor grid in the pentode is omitted for clearness

equivalent to grid tick in a transmitter (see page 6) and is due to the same causes.

It may be avoided by proper choice of condenser-leak time constant, and by proportioning the reaction winding to give higher feed-back at lower frequencies, reducing the general level of feed-back at the same time. If trouble still persists a grid stopper may be introduced in series with the grid. This cuts down the voltage reaching the grid in proportion to the relative impedances of the stopper and the grid capacitance. Since the reactance of the latter falls as the frequency rises the grid voltage is automatically reduced at the higher frequencies. The value of grid stopper depends on the frequencies under consideration, and may easily be calculated from normal circuit laws.

## Mixing Technique.

The exact manner in which the beating or mixing of the

oscillations is produced has evolved through several stages. Early procedure used a simple heterodyne mixing operating with a square law detector. The local oscillation was strong and the incoming signal caused variations in amplitude, according to the usual heterodyne detection principle (see Vol. I, Chapter XIV). A typical circuit is shown in Fig. 40 in which a pentode valve, designed to have a square

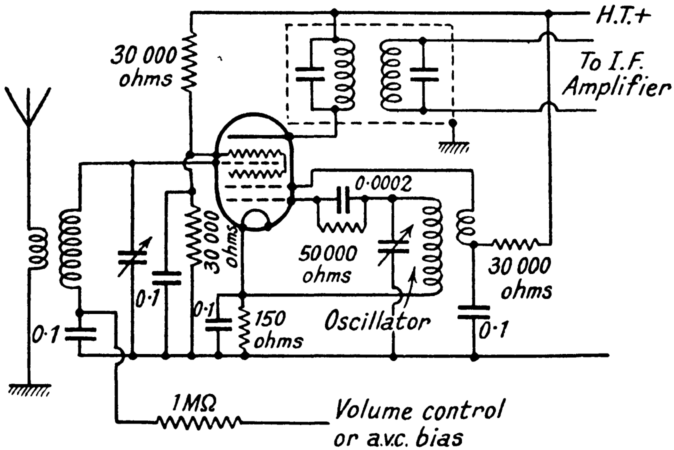



FIG. 41. TYPICAL PENTAGRID FREQUENCY-CHANGING CIRCUIT

law characteristic, was used as an anode bend detector. A small triode, assembled round the same cathode, generated the local oscillation, and the two were mixed by using a common bias resistor in the cathode circuit.

A later technique introduced a principle known as *electronic mixing*. The first valve of this class was known as the pentagrid or heptode, and was used as illustrated in Fig. 41. Next to the cathode were two grids which were used as grid and anode for the oscillator. The electrons flowing past the two grids to the valve proper were controlled very largely by the voltage on the first grid which was fluctuating at the frequency of the local oscillation. The incoming signal modulated this already fluctuating

electron stream and produced a composite anode current varying in the manner required.

To provide the outer section of the valve with tetrode characteristics, a screen was introduced between anode and signal grid while an additional grid was interposed between the signal grid and the oscillator section in order to screen one from the other, and avoid interaction between the two, other than due to purely electronic action. This grid was connected to the normal screen.

### Pulling.

This question of interaction is important, for if two circuits are coupled together they act as a combined circuit exhibiting two tuning points. At the high frequencies involved, the inter-electrode capacity inside the valve is sufficient to cause appreciable coupling, so that the tuning of the signal input circuit has an appreciable effect on the oscillator frequency. At the frequencies employed in short wave reception this pulling becomes very marked and is sufficient to upset the operation of the receiver quite seriously.

A further disadvantage is that the mixing is adversely affected. It is customary, for convenience of ganging as explained later, to use an oscillator frequency higher than the signal frequency. Under such conditions the interaction produced by inter-electrode capacitance coupling is in opposition to that produced electronically so that as the frequency rises the effectiveness of the valve as a mixer is seriously reduced.

Still another disadvantage of the heptode is that since the grid closest to the cathode has the greatest effect on the anode current, the oscillator grid variation is subjected to the full amplification of the valve, whereas the signal grid is only subject to a much smaller amplification. This is the wrong way round for efficient operation, since it is easy to generate a large local oscillation, whereas the signal being received is normally very minute.

### Triode-hexode.

These disadvantages led to the introduction of the hexode

mixing valve and its associated valve, the triode-hexode. In the hexode the signal grid is placed next to the cathode. Three further grids follow. The second and fourth are tied together and connected to an intermediate positive potential similar to the screen of a normal tetrode. The third grid in between these two is fed with voltage from the local oscillator.

The action of this valve is similar to the heptode except that the conditions are reversed. The incoming signal

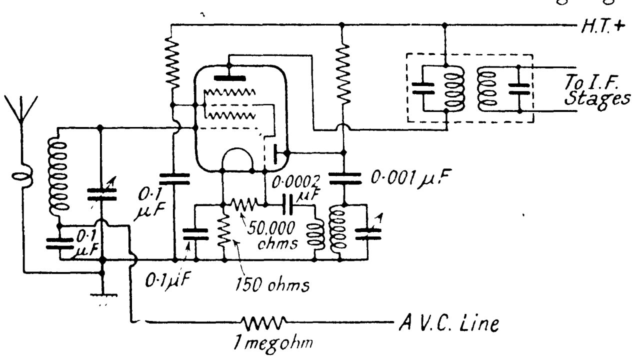



FIG. 42. TRIODE-HEXODE FREQUENCY CHANGER

causes variations in anode current which, because the valve is a high gain valve, are capable of causing appreciable anode current changes. These anode current changes are then modulated by the local oscillation applied to the modulator grid, and although this requires from 5 to 15 volts in order to produce a satisfactory performance, this is easily provided by a local oscillation.

In the triode-hexode a small triode section is mounted on the same cathode and the grid of this section, while available externally, is also connected internally to the modulator grid. A triode-hexode circuit is shown in Fig. 42.

The advantage of this type of valve is partly a better signal-to-noise ratio resulting from the amplification of the incoming signal before it is modulated by the local oscillation, and secondly appreciably less interaction between signal and oscillator circuit. This is further enhanced by

using a tuned anode circuit in the oscillator, so that the circuit which actually controls the frequency of the oscillation is considerably removed from the signal circuit. A tuned anode circuit is also more stable, being less dependent in frequency on circuit changes than a tuned grid arrangement.

## Theory of Mixing.

The operation of the electronic frequency changer is essentially different from the earlier form of rectification process. Where a detector is used to produce the mixing, the local oscillation and the incoming signal are *added.* The sum of two sine waves produces a composite wave having frequency equal to the mean of the two component waves modulated at a frequency of half the difference. This modulation component is simply a variation of amplitude, and cannot be detected unless rectification is present, which process incidentally produces a doubling term which gives the beat tone equal to the actual difference between the frequencies. Since one of the inputs is very small, it is customary to adopt a square law detector, in which case the familiar heterodyne amplification is obtained.

Let $v_1 \sin \omega t$ and $v_2 \sin (\omega + f)t$ be the signal and local oscillations respectively. Using a square law detector, the anode current will be proportional to

$$(v_1 \sin \omega t + v_2 \sin (\omega + f)t)^2$$
$$= v_1^2 \sin^2 \omega t + v_2^2 \sin^2 (\omega + f)t + 2v_1v_2 \sin \omega t \sin(\omega + f)t$$

The last term is the one in which we are interested, and this may be rewritten in a more interesting form. By the usual trigonometrical rules

$$2v_1v_2 \sin \omega t \sin (\omega + f)t = v_1v_2 (\cos ft - \cos(2\omega + f)t)$$

The second term here denotes a frequency of approximately twice that of the incoming signal, but the first term is a low-frequency term representing an oscillation at the difference frequency $f$, and having an amplitude proportional to both $v_1$, the incoming signal, *and* $v_2$, the local oscillation.

With electronic mixing, the incoming signal is made to

modulate the local oscillation and the anode current is proportional to the *product* of the two component frequencies. In such circumstances two distinct frequencies are produced, one equal to the difference in frequency, and the other to the sum, the latter, of course, being disregarded and attention concentrated on the difference frequency.

Assuming inputs of $v_1 \sin \omega t$ and $v_2 \sin (\omega + f)t$ as before, the anode voltage will be proportional to

$$v_1 v_2 \sin \omega t \sin(\omega + f)t$$

$$= \tfrac{1}{2} v_1 v_2 (\cos ft - \cos(2\omega + f)t)$$

which contains a radio-frequency term and a difference term as before.

It is, therefore, not necessary to have any rectification present, and this enables the valve to operate under more efficient conditions. It is important to arrange the circuit so that it by-passes the sum frequency satisfactorily. This is usually achieved by designing the i.f. transformer in the anode of the frequency changer to operate with a tuning capacity of at least $50 \mu\mu$F., or preferably rather more, so that the anode impedance at the frequency $2\omega + f$ is very small, and the voltage developed is negligible, whereas at the difference frequency the anode impedance is high since the i.f. transformer is tuned to this frequency, and an appreciable voltage is developed.

It is worth noting that the expression for the difference voltage contains a 2 in the denominator, indicating that the output is one-half that which would be obtained by the valve operating under the same conditions on a single frequency. This is found to be the case in practice, and if a frequency changer is supplied with i.f. on the signal grid, the gain will be found to be twice as great as that which is obtained when the valve is operating as a frequency changer.

## Conversion Conductance.

In frequency changers a term analogous to the customary mutual conductance is usually quoted. This is the change

in anode current *at beat frequency* divided by the change in grid voltage at signal frequency, and it is known as the *conversion conductance*.

The gain of a frequency changer stage may be determined in the same way as an ordinary h.f. pentode, being equal to $cZ$, where $c$ is the conversion conductance and $Z$ is the anode impedance to the intermediate frequency. This, of course, assumes that the anode impedance is small compared with the internal resistance of the valve, which is justifiable since the valve is made with characteristics similar to those of an ordinary h.f. pentode.

While the mutual conductance of a h.f. pentode is usually of the order of three (and for special types of valve may be more than twice this figure) the conversion conductance of a modern frequency changer is less than one, usually being of the order of 0·5, or even less.

## Noise in Frequency Changers.

Shot noise is more severe in a frequency changer than with a normal valve, for various reasons. Referring again to the basic expression $v_s^2 = 2keF\Delta f i_a/g^2$ quoted on page 68, it will be clear that the factor $i_a/g^2$ is two or three times higher than with a normal valve while the factor $F$ is found to be of the order 0·5 to 0·6.

Up to a point the designer still aims to keep $g^2/i_a$ high, but this involves long cathodes and consequent increase in inter-electrode capacitances, which causes increased "pulling" between the circuits, and often a valve is deliberately designed to have more shot noise in order to permit improvement in these other respects.

Moreover, if the valve is fed with an amplified signal from one or more h.f. stages, the signal-to-noise ratio of the frequency changer becomes of minor importance, the Johnson noise on the input to the h.f. stage becoming the predominating factor.

It is sometimes stated that the presence of the local oscillation in a frequency changer causes the admittedly higher background noise which is observed. This is quite erroneous.

## Undesired Response.

Probably the greatest defect of the superheterodyne system is its liability to produce false signals or interferences. This arises from the large number of combinations which will affect the frequency changer. The most obvious form of false response is that known as *image* or *second channel* interference.

The local oscillator is adjusted to differ from the wanted signal by the intermediate frequency and is usually arranged higher than the signal frequency for convenience of ganging. It could, however, equally well be lower and it follows therefore that for any setting of the oscillator frequency there will be two signal frequencies, one lower and the other higher, each of which will produce the required intermediate frequency. Only one of these is the wanted signal.

To avoid this interference the pre-selecting circuit, i.e. the tuning circuit prior to the frequency changer, must be capable of discriminating between the wanted signal and another signal separated by twice the intermediate frequency, *even though this second signal may be many hundred times stronger*. With low intermediate frequencies this is not always easy. The criterion is actually the relative value of the intermediate frequency and the signal frequency, and as the signal frequency becomes higher it becomes desirable to increase the intermediate frequency correspondingly. The subject is discussed further in Chapter XII.

Broadcast receivers usually use at least two tuned circuits prior to the frequency changer, often associated with a high-frequency amplifying valve to improve the signal-to-noise ratio. This is adequate to avoid second channel interference on medium and long waves but is not sufficient for short waves where the interference is present, and each station will be found to show two tuning points.

## Whistles.

An even more annoying form of interference is that which gives rise to whistles at certain parts of the tuning scale.

These whistles arise from interaction between the oscillator or one of its harmonics and some received oscillation which differs by the i.f. plus or minus some small amount. This oscillation will be accepted by the i.f. amplifier and since it is also handling the carrier of the wanted station (converted to intermediate frequency) the two oscillations will beat and produce an audible whistle.

Another form of interference is obtained if two strong local stations differ in frequency by an amount equal to the intermediate frequency chosen. These will beat to produce an oscillation of the order of the i.f. which in turn will beat with the carrier of the station being received and will produce an audible beat.

Ideally, of course, no signal should reach the frequency changer at any frequency other than that required, but in practice the pre-selection is by no means perfect, and strong local stations will invariably give rise to these beats. The only remedy is to improve the quality of the tuning prior to the frequency changer so that the residual voltage which still comes through the chain when the circuit is off tune is very small. With receivers working on a fixed frequency, or within a restricted band, it is possible to choose the intermediate frequency so that the beats which would be produced are outside the audible range. A simple calculation of frequency differences, taking into account not only fundamentals but harmonics up to the fifth, will soon show what interferences are likely to occur.

With a broadcast receiver, owing to the wide range and the different conditions under which it has to be used so that the "local" stations are different at different parts of the country, it is not practicable to make any very definite choice and adequate pre-selection is the only remedy.

Another form of whistle arises from harmonics of the intermediate frequency. If reception is attempted at a frequency corresponding to a harmonic of the i.f. and if any coupling exists between the second detector and the input of the set, then whistles can arise since the second detector is bound to produce some measure of harmonics of the intermediate frequency.

In any case very careful screening and layout is essential

on the stages prior to the mixer. Otherwise, signals will find their way on to the signal grid of the frequency changer

(*a*) by direct pick-up on the signal frequency circuit of the frequency changer (or the oscillator circuit);

(*b*) by leakage from the aerial circuit to the frequency changer via stray capacitances or conductive paths which will short-circuit the selecting action of the tuned circuits.

The choice of earthing points is important, for a relatively short length of common earth lead will form a direct coupling from the beginning to end of the chain and seriously reduce the selectivity, as explained on page 65.

## Ganging in Superheterodyne Receivers.

The usual practice is to tune all the signal-frequency circuits, including the oscillator, together. The ganging of the band-pass and/or h.f. stages is simple and is effected by the methods described on page 62.

The oscillator circuit, however, is more troublesome owing to the entirely different frequency range to be covered. For example, using a 450 kc/s. i.f. the medium-wave signal frequency range will be from, say, 550 to 1 500 kc/s., while the oscillator must range from 1 000 to 1 950 kc/s. Reducing the inductance of the oscillator coil will help matters, but we are still left with the fact that the capacitance range of the condenser will be too great.

Thus to cover a range of 550 to 1 500 kc/s. a capacitance range of $(1\,500/550)^2 = 7{\cdot}45$ will be required. If we make the oscillator inductance such that it will tune with the minimum capacitance to 1 950 kc/s., it will give 715 kc/s. at the maximum, whereas we require 1 000 kc/s. Clearly the capacitance range must be restricted.

This can be done by including a series or "padding" condenser in the oscillator section of the gang condenser which will reduce the effective maximum capacitance to a value sufficient to tune the particular oscillator inductance to 1 000 kc/s. Such a capacitance will be fairly large and will thus have little effect on the minimum, but we have

no guarantee that the *tracking* will be correct at intermediate points on the tuning scale.

An alternative method is to add a large "trimming" condenser in parallel with the oscillator section. This again restricts the capacitance range and a value can be so chosen as to give correct tracking at top and bottom of the scale (though with a different oscillator inductance).

As before, the intermediate tracking may be incorrect and actually both methods produce errors in opposite

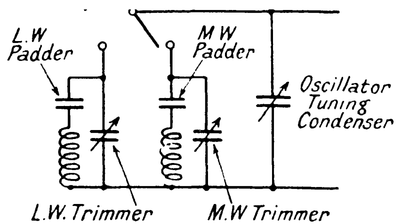



FIG. 43. TRACKING CIRCUIT FOR SUPERHETERODYNE OSCILLATOR

directions. Hence, in practice, a combination of the two is used, a third point being chosen in the middle of the scale at which it is arbitrarily assumed that the tracking shall be correct. We then have three "spot-on" frequencies and three variables—the oscillator inductance, the padder and the trimmer, and we can find definite values for each which will fulfil the three conditions. Elsewhere the tracking is still a little out, but the method can be made to give very satisfactory results.

Alternatively, specially shaped vanes may be used on the oscillator section of the condenser, which is made to have a smaller maximum capacitance than the other sections, but such a condenser can only be used with a specified i.f. and on one wave-band. For any other wave range padders and trimmers must still be used.

The padding and trimming capacitances must, of course, be altered for each wave range and are usually incorporated in the coil system and changed over by the switching which alters the coils. One arrangement is shown in Fig. 43.

**Frequency Drift.**

A most important practical consideration is that of frequency drift. As the parts of the receiver warm up the oscillator frequency tends to change and with bad design this drift is serious, necessitating constant re-tuning. Moreover the drift may continue almost unabated for some hours.

John M. Miller has shown (*Electronics*, Nov. 1937) that this is due to serious change in the dielectric constant of some materials used for insulation of the tuning condenser. Bakelized linen, for example, shows a change of + 0·2 per cent per degree C. Ebonite and high grade synthetic compounds are subject to changes of the order of + 0·02 per cent.

Ordinary ceramic materials have a positive temperature coefficient between 0·01 and 0·02 per cent, but certain materials incorporating oxides of titanium have a negative coefficient of 0·06 to 0·07 per cent. By allocating a suitable proportion of the circuit capacitance to this negative-coefficient material the positive coefficient not only of the remaining capacitance but also the inductance in the circuit may be offset, so that the circuit is substantially free from frequency drift.

**Flutter.**

Frequency flutter is an effect arising if the a.v.c. control voltage causes a change in oscillator frequency. Then if the frequency drifts the detector voltage falls causing alteration of the bias on the frequency changer (via the a.v.c.) which pulls the frequency back. The detector voltage rises again, removing the a.v.c. and the original state is resumed. The detector voltage is thus in a continual state of fluctuation causing a disagreeable fluttering of the speech or music.

There are various other points of design which could be considered but since this book is concerned with fundamentals they cannot be dealt with in detail.

## EXAMPLES V

(1) The i.f. transformer of a superhet comprises two circuits each of 10 000 $\mu$H. inductance tuned with a capacitance

of $150\,\mu\mu F$. The circuits are critically coupled and each has an effective resistance of 150 ohms. ($f = 130$ kc/s.) Calculate—

(*a*) The mutual inductance for critical coupling (use the criterion $M^2\omega^2 = R_1R_2$).

(*b*) The effective primary resistance and the "dynamic" impedance $L/CR$ of the primary.

(*c*) The approximate gain with a screen-grid valve having a slope of 2 mA./V.

(*d*) The proportion of the primary voltage developed across the secondary.

(2) If the transformer of the preceding question is used in the anode circuit of a frequency changer having a conversion conductance of 1 mA./V., what will be the overall gain (signal voltage input to i.f. voltage across the secondary of the transformer)?

# CHAPTER VI

## THE RADIO RECEIVER

### (c) The Detector Stage

A very important link in any radio receiver is the detector. This should normally obey a linear law—i.e. the rectified (l.f.) output should be proportional to the h.f. input. This is particularly important for telephony reception, not only on the score of freedom from distortion (so that the l.f. output may follow faithfully any variations in the h.f. input due to modulation) but also from considerations of selectivity.

It has been found that with a linear detector any signal which is stronger than another (interfering) signal will take control. Although the rectified output will contain a radio frequency made up of a combination of both the wanted and the unwanted station, only the modulation of the stronger signal will be heard.

The ultimate aim of the tuning circuits is thus to reduce any station to which the circuits are not tuned below the level of the wanted carrier (to which the circuits *are* tuned). For further information the reader is referred to a paper by Beatty on "The Apparent Demodulation of a Weak Signal by a Stronger One," *Wireless Engineer*, Vol. 5, page 300, June, 1928.

### Diode Detectors.

The simplest form of detector is the diode, which is practically linear and has a high rectification efficiency. Fig. 44 shows a simple diode circuit. The h.f. voltage across the tuned circuit will produce a current through the diode which will charge the condenser to a voltage approximately equal to the peak value of the applied e.m.f. During the remainder of the cycle this charge will leak away through the resistance $R$ until the next peak, when the diode will again conduct and the charge will be restored.

If the signal is modulated, the time constant of the resistance and condenser must be such as to allow the charge on the condenser to vary at least as rapidly as the highest modulation frequency. This involves relatively low values of $R$ and $C$, but the leak cannot be made smaller than about 0·25 megohm or the damping on the tuned circuit becomes serious, while the condenser must be large compared with the self-capacitance of the diode or an appreciable fraction of the voltage is lost. A value of about 100 $\mu\mu$F. is usual.

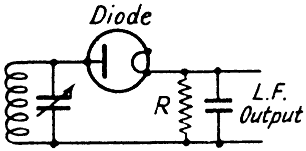



FIG. 44. SIMPLE DIODE DETECTOR

Fig. 45 shows a slightly modified form of circuit in which the leak is across the diode instead of across the condenser. The action is similar and it is desirable to analyse the

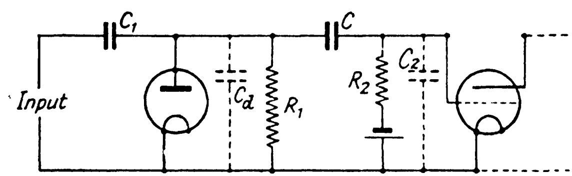



FIG. 45. PARALLEL DIODE CIRCUIT

operation in greater detail. It will be seen that the diode is coupled to the succeeding valve through an isolating condenser (to insulate the grid of the second valve from the steady potential developed across the diode resistance $R_1$). A grid leak taken to a suitable source of bias is connected across the output in the usual way.

The application of a modulated carrier wave across the input charges the condenser $C_1$ to a voltage approximately equal to the instantaneous peak value of the applied e.m.f. Since the carrier is modulated this condenser charge will be in the form of a steady voltage with a superposed alternating component.

It will be clear that the circuit offers different impedances

to the two types of voltage. To the steady d.c., the impedance is simply $R_1$. The alternating component, however, has a parallel path through $C$ and $R_2$ while in addition the shunt capacitances $C_1$, $C_2$ and $C_a$ also bypass some of the current.

The alternating component therefore will produce relatively more current than the d.c. component, so that distortion is liable to be introduced. The maximum modulation which can be handled without distortion can be arrived at as follows.

Let the mean value of the condenser voltage be $e$. Then the modulation component will be $me \sin pt$ where $m$ is the depth of modulation and $p$ is the modulation frequency. If $Z$ is the effective impedance of the circuit to a.c., the modulation current will be $(me/Z) \sin pt$, and the peak value of this will be $me/Z$.

Now this peak value cannot exceed the d.c. current. Hence we can write $me/Z = e/R_1$, whence

$$m = Z/R_1$$

At low frequencies the capacitances exercise a negligible effect so that $Z$ is simply $R_1R_2/(R_1 + R_2)$. The maximum modulation depth is thus $m = R_2/(R_1 + R_2)$.

At higher frequencies $Z$ is less than the value just quoted so that the permissible modulation falls off still more. By making $R_2$ three or four times $R_1$ we can approach full modulation, but some distortion is unavoidable. Special networks have been suggested to obviate the difficulty, but the usual practice is to make $R_2$ large and admit a small distortion on the peaks.*

It is worth noting that up to the critical value the efficiency of rectification is very high, because although the a.c. impedance is lower than $R_1$ the current is greater, and in fact the a.c. modulation voltage is the same as it would be if the circuit only contained the one resistance $R_1$. The rectification efficiency is then $R_1/(R_1 + r)$, where $r$ is the diode resistance, this being the ratio of the a.c. volts across $R_1$ to the modulation voltage across the whole circuit.

Diode characteristics of the type shown in Fig. 46 are

* See *Wireless World*, Dec. 28th, 1934.

often used. These show the d.c. through the diode for different values of steady negative bias with various values of applied a.c. We can draw a load line $OPA$ through $O$ corresponding to the d.c. load $R_1$. Where this cuts the appropriate input curve will show the operating point $P$

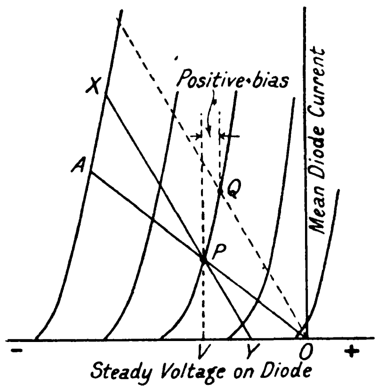



FIG. 46. DYNAMIC DIODE CHARACTERISTICS

The various curves are for different values of applied a.c.

and the steady bias $OV$ which will be developed across $R_1$. If through $P$ we draw another load line corresponding to the a.c. impedance $Z$ we can see at once that the excursion is limited since $XPY$ cuts the zero axis before full modulation is developed.

A possible remedy is to apply positive bias to the diode which would shift the operating point to the right as shown dotted. With this arrangement (first suggested by Kirke) the distortion may be avoided, but a permanent diode current then flows, giving heavy damping on the preceding circuit.

With suitable precautions the diode can be made to give particularly faithful detection, and this, coupled with its ability to handle large inputs of the order of 10 to 20 volts has made it very popular in recent years. Many circuits to-day use a diode to drive the output valve direct, without any intervening stage. Special high sensitivity

valves are used capable of delivering the full ($2\frac{1}{2}$ watts) output with an input of only $2\frac{1}{2}$ volts r.m.s.

Special double-diode and double-diode-triode valves have come to the fore recently, particularly for a.v.c. circuits, as explained later.

## Triode Detectors—The Anode-bend Rectifier.

The use of a triode as a detector has been discussed in Volume I. There are two principal methods of use. In the first the valve is used as an anode-bend arrangement in which the curvature of the characteristics at large values of grid bias is utilized. A decrease in grid voltage then produces an appreciable change in anode current, whereas an increase makes little difference as the current is already practically zero.

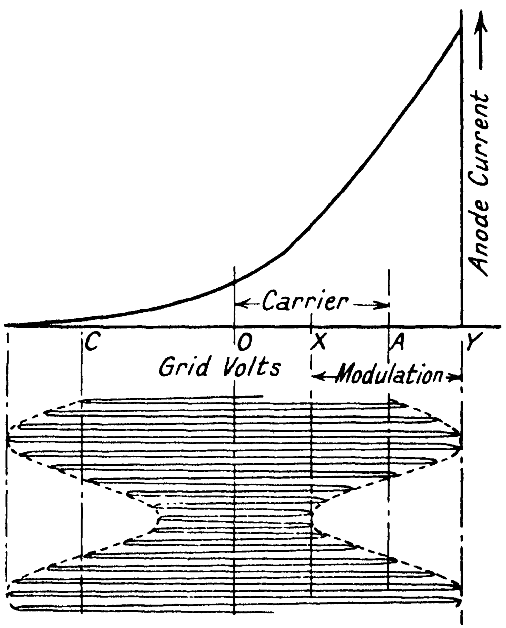



FIG. 47. USE OF ANODE BEND DETECTOR DISTORTIONLESS CONDITION

This method of rectification is only used under special circumstances. It is, of course, insensitive because of the very low slope on the characteristic at the point where the anode current is nearly zero. Secondly, the slope increases as the grid swing increases. This can clearly be seen from Fig. 47, and means that the rectified current is not strictly proportional to the applied signal. Hence, where distortionless operation is required, this form of detector is not suitable, except under certain limited conditions.

A telephony transmitter is rarely modulated more than 80 per cent. If the valve is designed to operate with a high anode voltage and a large negative bias, the curved portion at the bottom of the characteristic only comes into play with small input. Suppose the carrier input is such as to swing the valve over the region $OAC$. Variations in carrier strength due to modulation would vary the amplitude between the limits $X$ and $Y$ which are still on the relatively straight part of the curve. Under these conditions distortionless detection would result, but it will be seen to require a large carrier signal and, moreover, the grid bias should be adjusted exactly to suit the incoming signal.

## Triode Detectors—The Grid Rectifier.

In the second method, the audio-frequency voltage developed at the grid by the rectifying action causes the anode current to vary so that considerable amplification is obtained. The valve, in fact, operates exactly as a diode coupled to a simple l.f. amplifier, but the presence of high-frequency voltages on the grid disturbs the operation. These h.f. voltages are considerably greater than the low-frequency modulation, and therefore they overload the valve long before full l.f. output has been obtained. Hence, the input which the grid detector will handle is limited.

The action may be considered briefly as follows: The application of a signal to the grid will build up a voltage on the grid condenser due to the action of the grid as a diode. This voltage will be approximately equal to the peak value of the signal, i.e. $(\sqrt{2})v_g$, where $v_g$ is the r.m.s. carrier voltage. Superposed on this will be a modulation

voltage having a peak value $m$ times the steady voltage, where $m$ is the modulation depth. The anode voltage will he obtained by the usual formula for the amplification of a triode, using the effective $\mu$ and $r_a$ for a value of bias corresponding to the voltage to which the grid condenser has charged.

As the input signal increases the bias becomes greater, and at a comparatively small input the working point on

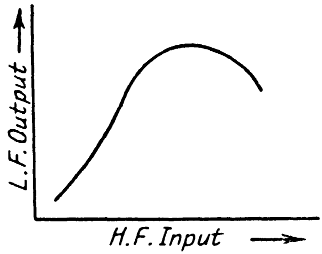



FIG. 48. ILLUSTRATING LIMITING WITH GRID DETECTOR

the characteristic runs off the linear portion. When this happens, anode-bend rectification begins to occur, which acts in the opposite phase, and the effective anode voltage is due to the sum of the two effects acting simultaneously. The result is a limiting action as shown in Fig. 48.

Terman has devised a simple treatment of the case. As explained later, the anode impedance to radio frequency is made low to reduce Miller effect. This in itself reduces the permissible anode swing, and it will be found that the correct grid bias for zero load is appreciably less than the value for the same valve as an amplifier with normal load. Let this "no-load" bias be $k$ times the normal value.

If $E$ is the peak value of the carrier the peak modulated signal across the grid condenser is $E(1 + m)\beta$, where $m$ is the depth of modulation and $\beta$ is the rectification efficiency $= R_1/(R_1 + r)$. The most negative signal on the grid will be this value of grid condenser voltage *plus* the maximum signal input $E(1 + m)$, i.e. $E(1 + m)(1 + \beta)$. For distortionless working this must not exceed twice the normal bias $V_g$, which, as we have seen, is $kV_0$, where $V_0$ is the correct bias for the valve acting as an amplifier.

The maximum permissible peak input to the valve as an amplifier is $V_0$ (equal to the grid bias). Hence we can write

$$\frac{\text{Maximum carrier input}}{\text{Maximum amplifier input}} = \frac{k}{1+\beta} \text{ for 100 per cent modulation.}$$

Since $k$ is usually about 0·7 and $\beta$ is about 0·8, this expression is of the order of 0·4, which explains the ease with which a grid detector will overload.

Quality with a properly adjusted grid detector can be quite good with small signals. The conditions are somewhat similar to those of a positively-biased diode since the grid resistance is usually returned to a small positive voltage. (In an indirectly heated valve it is taken to cathode, but since grid current starts to flow at about $-$ 1 volt the effect is the same.) On the other hand the d.c. and a.c. circuits are now the same apart from the effect of shunt capacitances. Generally speaking, therefore, a grid detector will handle a higher depth of modulation than the usual diode circuit will permit.

The efficiency of rectification with a grid rectifier, however, is usually lower than with a diode, because the a.c. grid-filament resistance $r$ is higher due to the open spacing of the grid and the presence of positive potential on the anode which attracts many electrons which would otherwise stop at the grid.

## Miller Effect—Input Admittance.

The feed-back of energy from anode to grid of a valve was first analysed by John M. Miller,* who obtained expressions for the input admittance of a valve. (Admittance is the reciprocal of impedance.)

Fig. 49 shows the equivalent circuit of a valve, this being a simple network except for the fact that there is the amplified voltage $\mu e_g$ in the anode circuit. By ordinary a.c. theory it can be proved that

$$\text{Input resistance } R_g = -\frac{1/\omega C_{ga}}{A \sin\theta}$$

* *Bureau of Standards Bulletin, No.* 351.

and input capacitance $C_g = C_{gf} + C_{ga}(1 + A \cos\theta)$
where $A$ is the effective amplification of the valve

$$= [Z/(r + jZ)]\mu$$

and $\theta$ is the angle by which the voltage on the load leads the a.c. anode voltage $\mu e_g$.

Let us examine these equations further. The input capacitance is obviously always greater than the "geometric" capacitance of the valve, $C_{gf}$, while the input resistance

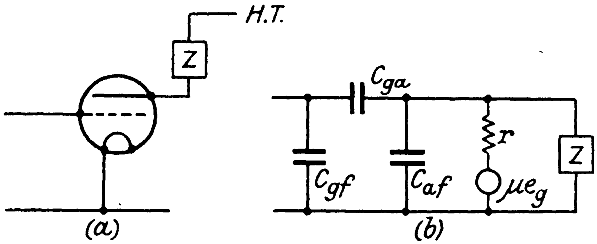



FIG. 49. EQUIVALENT CIRCUIT OF TRIODE

may be positive or negative according to the sign of the denominator $A \sin\theta$. If it is positive it means that extra damping is introduced into the circuit. If negative, the damping is reduced and self-oscillation may occur.

If the anode load is inductive (at the frequency under consideration), $\theta$ is positive and hence the input resistance is negative. This is the cause of instability in h.f. amplifiers. Cos $\theta$ is fairly small and $C_g$ is equal to $C_{gf}$ plus two or three times $C_{ga}$.

If the anode load is resistive, $\theta$ is zero and the input resistance is infinite. This condition can only be obtained by a tuned anode circuit, and hence a h.f. stage with the circuits dead in tune is not unstable at that frequency (but it *is* at lower frequencies, which make the anode load inductive). $C_g$ would be very high since $\cos\theta = 1$, but the condition is an impracticable one and need not be analysed further.

With a capacitative anode load $\sin\theta$ is negative and the input resistance is positive, i.e. the damping is increased.

Cos $\theta$ remains positive (for $\theta$ never exceeds $-(\pi/2)$) and $C_g$ is thus always more than $C_{gf}$, the limiting value being $C_{gf} + C_{ga}$ when cos $\theta$ is nearly zero.

**Bypass Condenser.**

The application of these formulae to h.f. circuits was discussed in Chapter IV. It is interesting, however, to consider the effect in a detector circuit.

The anode impedance of a detector is either a resistance or a (low-frequency) inductance, both of which appear capacitative at radio frequencies owing to self-capacitance. The effective amplification $A$ at high frequencies will depend on the reactance of this capacitance load relative to the valve, and if we are to keep the input resistance high the value of $A$ should be small. In other words, the capacitance should be large, and in practice it pays to connect a bypass condenser from anode to cathode to ensure that the valve shall have little amplification at radio frequencies.

This capacitance must not be large enough to exercise any serious shunting effect on the audio frequencies, and a value of 100 to 300 $\mu\mu$F. is usual.

**Input Resistance.**

The input resistance of a detector is often required to be known. Let us consider a diode first of all. Since it is only a unidirectional conductor, special methods have to be adopted to estimate the input resistance. A simple method, also due to Terman, is to assume that the diode only conducts on the peaks of the signal. The power absorbed is then $\hat{E}i_{mean}$ where $\hat{E}$ is the peak voltage and $i_{mean}$ the average diode (or grid) current.

$i_{mean} = \beta\, \hat{E}/R_1$ where $R_1$ is the diode (or grid leak) resistance and $\beta$ is the rectification efficiency.

Thus power absorbed $= \beta\hat{E}^2/R_1$

But $\hat{E}^2 = 2E_{r.m.s.}$ so that power $= \dfrac{E^2}{R_1/2\beta}$

Hence the effective input resistance $= R_1/2\beta$.

$\beta$ is usually about 0·7 to 0·8.

A grid rectifier can be dealt with in similar fashion, but an additional allowance must be made for the Miller effect. This can be assessed by using the formulae on page 98, though, if the detector is adequately bypassed, the Miller effect will be small and laborious (and possibly fruitless)

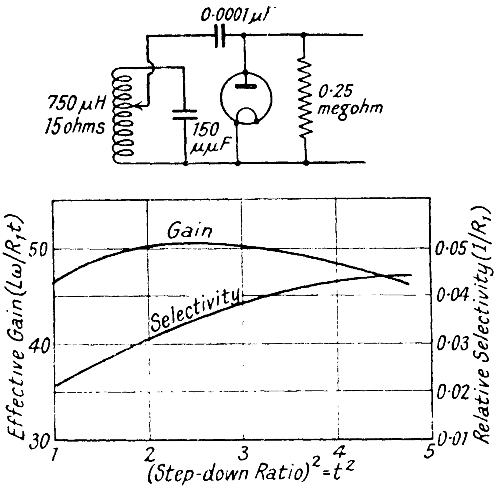



FIG. 50. SHOWING EFFECT OF TAPPING THE DETECTOR DOWN THE COIL

calculation can be avoided by adding about 20 per cent to the value already obtained.

With an anode-bend detector the Miller effect is small because of the very low slope of the valve at the operating point, unless the valve is being used as shown in Fig. 47. In any case, if the valve is adequately bypassed the input resistance should be very high, but if it has to be estimated the formulae already given may be applied.

## Design of Input Circuit.

If the detector load is serious, as it is with a diode or grid detector, it may be worth while to consider tapping down the input circuit.

It has already been shown (page 58) that the presence

of a resistance across a tuned circuit increases the effective series resistance. If the load is tapped across part of the circuit only, the extra damping is reduced by the square of the step-down ratio, while the voltage actually applied to the diode will only be reduced in direct proportion to the step-down ratio.

With a heavy load, therefore, the release of the damping may more than compensate for the reduction in voltage due to tapping down, and Fig. 50 shows a diode tapped across half the input circuit only.

The improvement in selectivity is often still more valuable and it may be that the voltage applied to the diode is no greater or even less than when the full circuit is used, but the tap is still worth while because of the markedly better selectivity.

The calculation of the optimum tap is simple. Let the load be tapped across a point such that the step-down ratio is $t$. Then the effective resistance across the whole circuit is $Pt^2$, where $P$ is the input resistance of the diode or other source of damping. This is in parallel with the dynamic impedance $Z = L/CR$ of the tuned circuit. Then

$$\begin{aligned}\frac{1}{Z_1} &= \frac{1}{Z} + \frac{1}{Pt^2}\\ &= \frac{CR}{L} + \frac{1}{Pt^2}\\ &= \frac{CRPt^2 + L}{LPt^2}\end{aligned}$$

But $Z_1 = L/CR_1$, where $R_1$ is the effective resistance

$$\therefore \frac{1}{Z_1} = \frac{CR_1}{L} = \frac{CRPt^2 + L}{LPt^2}$$

Whence $R_1 = R + L/CPt^2$.

It is not practicable to work out a formula for the gain in terms of $t^2$ as the expressions become unwieldy. In a practical case, knowing $L$, $C$, $P$ and $R$, it is easy to evaluate $R_1$ for various values of $t^2$, and then to plot the gain

$$= \frac{L\omega}{R_1 t} = \sqrt{\frac{L}{C}} \cdot \frac{1}{R_1 t}$$

against $t$ when the maximum will readily be found. The curve of Fig. 50 illustrates the type of variation obtained with a circuit having the constants shown. It will be seen that unless the load is very heavy or the circuit very good the increase in gain is small, the increased selectivity being the determining factor.

## Screen-grid Detectors.

The tetrode or screened pentode makes a good detector, owing to the high amplification factor and very small

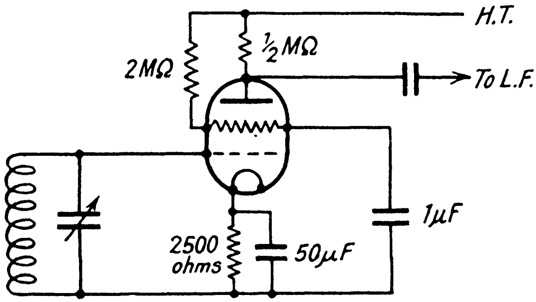



FIG. 51. SCREEN-GRID DETECTOR CIRCUIT

Miller effect, but it is essential to use resistance coupling owing to the high internal resistance of the valve. Otherwise a very large transformer would be required to provide a reactance comparable with that of the valve at anything like a low frequency. The usual resistance is of the order of 100 000 ohms with a grid detector, and 0·25 or 0·5 megohm with anode-bend detector.

With a.c. valves a self-bias arrangement can be used, and this automatically adjusts itself to a good rectifying condition. Satisfactory and reasonably distortionless rectification of high sensitivity can be obtained because the increase in anode current due to the rectifying action causes an increase in bias which limits the current. Hence, the usual square-law characteristic is not obtained, the output being practically linear as shown in Fig. 52.

Miller effect is, of course, practically non-existent owing to the screening in the valve, and hence little power is

drawn in the operation of the detector. An input voltage of 0·025 volts is capable of developing to 10 volts l.f. in the anode circuit with full modulation. The screen should be fed through a series resistance, as in Fig. 51, which enables the voltage to adjust itself to the optimum. With a potentiometer feed the setting is very critical.

## Automatic Volume Control.

It is convenient at this point to discuss methods of automatic volume control. Due to the fading which is

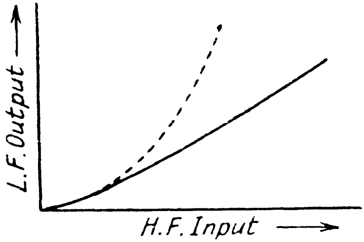



FIG. 52. ILLUSTRATING LINEAR ACTION OBTAINED WITH SELF-BIASED DETECTOR

experienced on foreign stations, the strength may vary from minute to minute in a disconcerting manner. This may be minimized by adjusting the amplification of the h.f. (or i.f.) stages in accordance with the voltage developed at the detector. The high-frequency carrier should, of course, remain constant, irrespective of the low-frequency modulation, and if it does not do so then some compensation may be provided.

Fig. 53 shows a typical diode a.v.c. circuit. Across the diode resistance is developed a steady voltage proportional to the mean value of the carrier, and low-frequency variations are superposed on this mean value. The l.f. voltages are transferred to the valves following the rectifier, and the d.c. potential is used to control the amplification of the h.f. stages which are provided with vari-mu valves.

If the carrier voltage at the detector increases, the steady voltage across the diode resistance also increases. This will be seen from the figure to provide a negative voltage at

the point $X$, and this point therefore is connected back to the grids of the h.f. valves. Hence, the grid bias on these valves increases, causing a reduced amplification, and this checks the increase in the carrier at the detector.

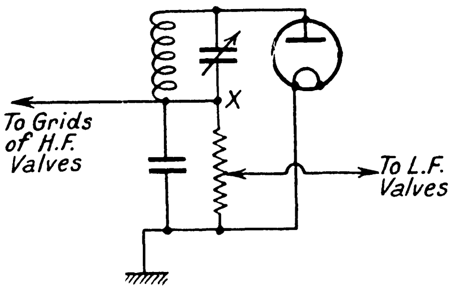



FIG. 53. DIODE A.V.C. CIRCUIT

A complete levelling out of the characteristic cannot be obtained because the detector voltage must increase slightly in order to produce the necessary control voltage. The

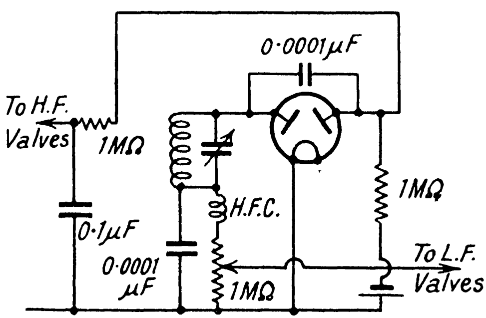



FIG. 54. CIRCUIT FOR PROVIDING DELAYED A.V.C.

larger the number h.f. or i.f. stages in front of the detector, however, the more sensitive is the control and with commercial receiving circuits a large number of stages is used for this reason.

Sometimes it is practicable to include a vari-mu low-frequency stage following the detector and to control this

as well, and if this is done variations in the input of over 10 000 to 1 will cause no appreciable change in the output. It must be remembered that this device only operates on the carrier and is neither affected by, nor does it affect, the low-frequency modulation.

## Delayed A.V.C.

In some cases it is desirable to arrange that the automatic volume control does not come into operation until a certain minimum strength of signal has been reached. This is accomplished by using two diodes, one of which provides the signal frequency voltages and the other, which is tied in parallel with the first, provides the a.v.c. voltage.

This second diode, however, is provided with a small permanent negative voltage so that rectification does not take place until the peak value of the applied signal has exceeded this delay voltage. Fig. 54 shows a circuit of this type. The operation is otherwise the same as already described. It should be noted that in order to filter out any radio-frequency variations from being fed back to the h.f. stages, a decoupling circuit is included consisting of a high resistance—usually one megohm—and a fairly large condenser of about 0·1 $\mu$F. or more. This has no effect on the d.c., but it filters out any radio-frequency currents.

It is important that the time constant of the a.v.c. and decoupling circuits should not be too large. Otherwise there is an appreciable time lag between the variations in the carrier and the correcting action. Sudden variations in strength, such as are obtained when tuning in, may momentarily paralyse the receiver because a high voltage is built up on the condensers which takes an appreciable time to leak away.

## Amplified A.V.C.

The aim of the designer is to obtain a good *control ratio*, i.e. a large control voltage for a small change in signal strength. To some extent, delayed a.v.c. automatically provides an improvement, as the curves in Fig. 55 show.

The full curve represents the output voltage in terms of input, without a.v.c. If plain a.v.c. is used, this becomes

operative at once and cannot therefore be allowed to be too fierce in its operation or the output would never reach its normal value. The chain dotted curve shows a typical performance without delay. It will be noted that the gain on weak signals is appreciably reduced.

The third curve shows delayed a.v.c. which does not operate until the output is nearly up to the full amount

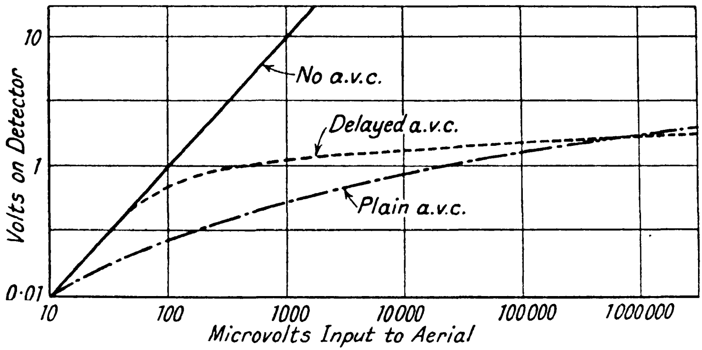



FIG. 55. SHOWING THE EFFECT OF AUTOMATIC VOLUME CONTROL ON RECEIVER OUTPUT

and then comes into play with full action, giving an almost level output and a much better control ratio.

It may be that the voltage at the detector which delivers full output is insufficient to give a sharp enough control. Thus with a delayed system requiring 5 volts radio frequency to load the detector, we should use, say, 4 volts delay. This gives only 1 volt to provide a.v.c. which is quite inadequate and actually on strong signals some overloading would be inevitable.

Various forms of amplified a.v.c. have therefore been devised, in which the a.v.c. voltage is applied to a d.c. amplifier which develops an adequate output. The systems are fairly simple if a separate valve can be used. In broadcast receivers this cannot usually be permitted and such single-valve systems as have been used have proved rather tricky.

A partial improvement can be obtained by feeding the

a.v.c. from the anode of the valve preceding the detector, whereas the signal diode comes off the secondary where, as explained in Chapter V, the voltage is reduced by one-half. The a.v.c. voltage developed is thus twice as great.

In commercial practice a separate i.f. amplifier is used for the a.v.c., supplied from some point subsequent to the frequency changer. This amplifier is *not controlled by the a.v.c.*, so that the full increase of signal is allowed to develop a.v.c. voltage. With the normal system, of course, as already explained, the action of the a.v.c. automatically cuts down the source of the a.v.c. voltage itself, so that it is impossible to obtain a truly level characteristic.

## Quiet A.V.C.

As a development of this system the circuit is sometimes arranged to give no audible output until the carrier reaches a predetermined value. One method is to over-bias the output valve and allow the a.v.c. circuit to "knock off" this bias when the carrier voltage exceeds the delay voltage. The system is sometimes called "squelching" or "muting."

Another method is to put negative bias on the signal diode so that it normally does not operate. The a.v.c. diode, however, is left operating normally with a small delay if necessary. On the arrival of a signal sufficiently strong to operate the a.v.c., the signal diode is released and the circuit functions normally.

The release of the signal diode (or the l.f. stages) may be accomplished by suitable circuit arrangements of the normal network, but it is more usual to employ a separate muting valve for the purpose. Fig. 56 shows such a circuit. $V_1$ is a double-diode-triode, the first diode being used for a.v.c. and therefore made negative to the cathode by the drop in the resistor $R_1$ which is providing bias for the triode section of the valve in the usual way.

The second (signal) diode is coupled through a 0·0001 $\mu$F. condenser to isolate it from the first as regards d.c. potential but the leak in this case is returned to full negative so that it is negative to the cathode of $V_1$ by the voltage drops on $R_1$ and $R_2$. $V_2$ is a high-current high-slope valve. As soon as the a.v.c. diode operates, negative potential

is applied to the grid of $V_2$. The anode current rapidly decreases, knocking off the bias on the signal diode which thus comes into operation.

There are numerous alternative muting circuits, the general principle being the same. In all cases there is a

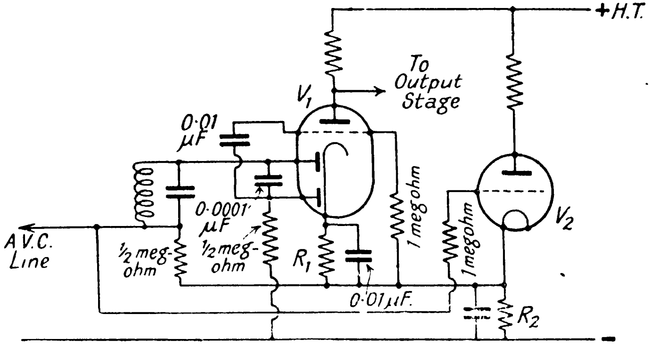



FIG. 56. MUTING CIRCUIT FOR INTERSTATION NOISE SUPPRESSION

point where the signal is only just able to release the squelch and distortion ensues. The aim of the designer is to make this transition very sharp so that no distortion occurs on any station having sufficient strength to be of programme value.

There are numerous variants of these basic principles in practical use to-day, but space does not permit a more detailed consideration of the subject.

## Tuning Indicators.

In a receiver fitted with automatic volume control, the strength of the signal does not vary appreciably as the station is tuned in. This is because the normal increase in strength as the tuning point is approached is counteracted by the automatic volume control coming into operation. It is, therefore, possible for an inexpert user to adjust the receiver quite incorrectly, to avoid which tuning indicators are often employed. These consist of a meter or other indicating device which is included in the anode

circuit of the controlled valves or otherwise connected thereto. As the bias on the h.f. valves is run back the anode current decreases, and at the point where the carrier is a maximum (i.e. at the tuning point) the anode current reaches its lowest value. Hence, a meter in the anode

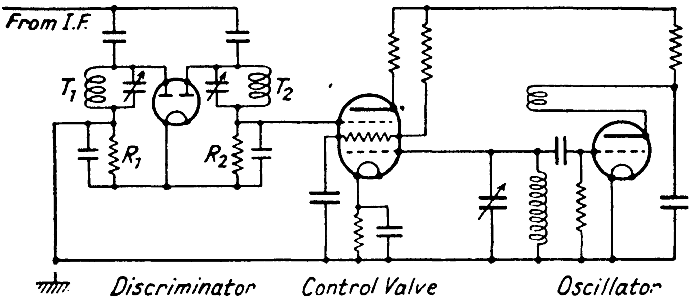



FIG. 57. DISCRIMINATOR CIRCUIT FOR OBTAINING AUTOMATIC FREQUENCY CONTROL

circuit would indicate quite definitely the exact tuning point, and devices of this nature are used quite appreciably.

## Automatic Frequency Control.

A development of some importance is the application of a.v.c. to control the *tuning* of a receiver. The system is principally applied to superheterodyne receivers where the fine control of the tuning is a function of one circuit—the oscillator. If the frequency of this circuit can be automatically adjusted to give correct tuning the effects of slightly incorrect setting and, more important, frequency drift can be largely counteracted. The method is known as a.f.c. (automatic frequency control) or a.t.c. (automatic tuning control).

This result is accomplished in the first place by using a discriminator circuit as shown on the left of Fig. 57. Two i.f. transformers are used, sharply tuned to peak a few kc./s above and below the true tuning point. If the oscillator setting is correct so that the intermediate frequency generated lies accurately in between these two tunes, both $T_1$ and $T_2$ will develop the same voltage, each being partially

off tune. The currents in the diode loads $R_1$ and $R_2$ will be equal and opposite and will offset one another, and no "control" voltage will be developed.

If the oscillator now drifts causing the i.f. to rise, the higher tuned transformer, say $T_1$, will develop more voltage, while $T_2$ will develop less. Consequently $R_1$ will develop more voltage than $R_2$ and there will be a resultant voltage applied to the control valve. If the oscillator drifts in the opposite direction, a similar but opposite control voltage will be produced.

Now this control voltage may be utilized in various ways. The most usual is to employ the Miller effect of a valve which depends on the effective amplification. In Fig. 57 the control is applied to the suppressor grid which is an effective method of altering the gain. A positive control voltage will cause the gain to increase, while a negative voltage will cause a decrease. The effective grid-cathode capacitance of the valve will change accordingly, and as this is in parallel with the oscillator tuning circuit the oscillator frequency will be modified in such a direction as to counteract the drift. The effective gain of the control valve must be arranged, by suitable choice of anode load, to produce this result.

An inverted Miller effect may be used, by inserting suitable phasing networks between anode, grid and cathode whereby the effective *anode* impedance may be made equivalent to an inductance or capacitance, dependent in either case on the gain of the valve. Still another method is to apply the control to the grid and rely on the changing a.c. resistance of the valve to damp the oscillator circuit to a variable extent, which again alters the frequency since the tune of a parallel resonant circuit is dependent on its effective resistance. For any of these last three arrangements the anode circuit of the valve is connected across the oscillator circuit.

There are various possible arrangements having different degrees of effectiveness, which need not be discussed in detail. The reader who desires further information should refer to an article entitled "A.T.C. systems," *Wireless World*, 19th Feb., 1937.

**L.F. Volume Control.**

The adjustment of the h.f. amplification prior to the detector has already been dealt with. It may still be, and often is, desirable to alter the subsequent amplification in the same way. This can conveniently be done immediately following the detector by using only a portion of the voltage developed across the detector resistance. The diode resistance in Fig. 53, for instance, is shown as a potentiometer, only a portion of the full low-frequency voltage developed across the resistance being applied to the succeeding stages.

Alternative positions for potentiometers are as grid leaks in resistance-coupled stages, or as a voltage-regulating arrangement connected across the secondary of a low-frequency transformer. The resistance element is usually made of composition where very high resistances are required, while for lower values up to about 100 000 ohms a wire-wound resistance track may be employed.

## EXAMPLES VI

(1) A signal of 4 volts h.f. is applied across the diode circuit of Fig. 43. Calculate the approximate voltage developed across the condenser.

If the signal is modulated with a steady note to a depth of 30 per cent, what will be the l.f. voltage component?

(2) The amplification factor and internal resistance of a grid detector are as shown below. Calculate the approximate l.f. voltage developed across an anode load of 20 000 ohms by a signal of 1 volt h.f. modulated 30 per cent.

| | | | |
|---|---|---|---|
| Grid bias . . . . | − 1 | − 1·5 | − 2 |
| $\mu$ . . . . . | 20 | 13 | 10 |
| $r_a$ . . . . . | 15 000 | 18 000 | 23 000 |

# CHAPTER VII

## THE RADIO RECEIVER

### (*d*) The Low-frequency Amplifier

The design of a satisfactory l.f. amplifier divides itself into three main headings, namely—

(1) Choice of suitable coupling impedances.
(2) Choice of suitable valve.
(3) Elimination of reaction effects.

The choice of coupling impedance depends upon the requirements. Where a single frequency is to be amplified, as in the case of certain telegraphic signals, it is possible to use coupling impedances having a high response over a narrow band of frequency, and even to use a tuned l.f. inductance or transformer.

For broadcasting or telephony reception, a range of frequencies has to be covered and the conditions are more stringent. Transformer coupling may still be used, the transformers become increasingly expensive as the frequency requirements become greater, but for high fidelity, resistance coupling is preferable.

For the amplification of frequencies above or below the usual musical range resistance coupling becomes essential. Direct coupling (referred to later) can also be employed, but it is neither so convenient nor so stable.

**Resistance Coupling.**

A resistance-coupled circuit is shown in Fig. 58, together with the equivalent circuit and vector diagram. The amplification is limited at low and high frequencies, the former by the effect of the coupling condenser $C$ and the latter by the stray capacitance $C_0$ shunted across the circuits and shown for convenience as all located across $R$.

It is possible to devise a mathematical expression for the gain taking both these factors into account, but it is

simpler and more conducive to clear understanding to separate the effects. We will first consider the gain at a middle frequency such that the reactance of $C_0$ is negligibly high by comparison with $R$ and the reactance of $C$ negligibly low compared with $R_1$.

Then the valve voltage $\mu e_g$ passes current through $r$ in series with $R$ and $R_1$ in parallel. The voltage transferred to the next stage is that across $R_1$, so that the gain is

$$P\mu/(P + r) \text{ where } P = RR_1/(R + R_1)$$

It must always be remembered that the presence of a high resistance in the anode circuit will reduce the anode

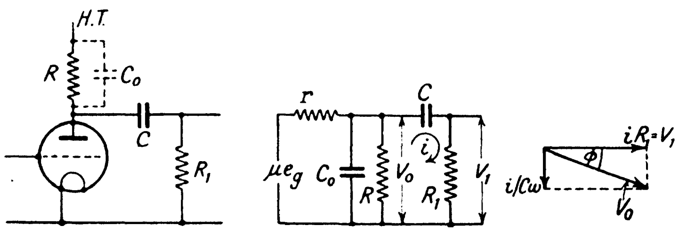



FIG. 58. SKELETON RESISTANCE-COUPLED STAGE WITH EQUIVALENT CIRCUIT

voltage, and if the h.t. supply has not been increased accordingly, $\mu$ may be much less than the nominal value.

In order to obtain maximum gain $R_1$ is usually made four or five times $R$, so that $P = R$ very nearly.

At low frequencies the reactance of $C$ is not negligible and the voltage $V$ is thus only a part of the voltage developed $V_0$. If $i$ is the current through $CR_1$, the proportion of $V_0$ developed across $R_1$ is

$$\frac{R_1}{(R_1 + 1/jC\omega)} = \frac{R_1}{\sqrt{(R_1{}^2 + 1/C^2\omega^2)}}$$

It is easy to calculate the relative values of $C$ and $R_1$ to make $V/V_0$ a given percentage, and a criterion often quoted is for the product $CR_1$ to be 0·0065 which makes $V/V_0 = 0{\cdot}9$ at 50 cycles. (It will be clear that the fundamental requirement is that the reactance of $C$ shall be low compared

with $R_1$ and so we can either increase $C$ or $R_1$. Thus it is the product of the two which is important.)

## Phase Angle.

This criterion is by no means adequate for high quality amplification, particularly where several stages are in use. The loss in each stage is cumulative, so that with three stages $V/V_0$ would only be 0·73.

A still more serious defect, however, is the change of phase which occurs, for the current $i$, and hence the voltage $V_1$ is by no means in phase with $V_0$ but leads by an angle $\tan^{-1}(1/RC\omega)$, as will be clear from Fig. 58. If $CR = 0{\cdot}0065$ this angle is 25·8 degrees, and three stages would give a lead of 77·4 degrees!

There are circuits where this is unimportant. It is stated by some authorities that the relative phase of the component frequencies in a musical transmission does not materially affect the result as heard by the ear, provided the amplitude of the components is correct, but this obviously must not be carried too far, while there are many circuits in which correct phase relationship is of vital importance. In such circumstances the product $CR$ must be increased to reduce the phase angle to the required value.

## High Frequency Loss.

At the upper frequencies the capacitance $C_0$ becomes troublesome and shunts $R$. The amplification is then reduced by the lowering of the effective impedance. $R$ and $C_0$ in parallel provide an impedance $Z = R/(1 + jRC_0\omega)$ and the gain is $\mu Z/(r + Z)$ taking due account of the vectorial nature of $Z$. Under given conditions of $C_0$ and $\omega$, the value of $R$ necessary to maintain, say, 90 per cent of the full amplification can be readily determined. [The higher $\omega$ or $C_0$ the less $R$ must be and this soon limits the effectiveness of the circuit.] $C_0$ includes the anode-cathode capacitance of the valve in question, the stray circuit capacitances, *and the input capacitance of the succeeding valve*. This latter capacitance may entirely swamp the previous ones for the input capacitance of a valve is not merely its static or geometric capacitance but is increased by the Miller effect.

## Miller Effect—Use of Screened Valves.

From the formula quoted in the last chapter the input capacitance $C_g = C_{gf} + C_{ga}(1 + A \cos\theta)$. Assuming an effective gain $A$ of 15 and $\cos\theta = 0{\cdot}8$, $C_{gf}$ and $C_{ga} = 7\,\mu\mu$F. each, we have $C_g = 7 + 7(1 + 12) = 98\mu\mu$F., a value which would seriously shunt any resistance above a few thousand ohms at 5 000 or more cycles per sec.

Hence, if effective amplification of the upper frequencies is to be maintained, $R$ must be kept low, which means a limited gain per stage. A remedy is to use screen-grid or pentode valves which have grid-anode capacitances of 0·001 $\mu$F. or less. Miller effect with such valves is negligible, but the static capacitance $C_{gf}$ still remains, plus the anode-cathode capacitance of the preceding valve and the stray capacitances.

The capacitance, indeed, is still of the order of 30 $\mu\mu$F. unless special valves are used so that the anode resistance cannot greatly exceed 100 000 ohms. Even so, however, the use of a screened valve is worth while, for the effective gain is roughly $gR$ (since $R$ is small compared with $r$ as explained in Chapter IV) so that with a $g$ of 2 under working conditions a stage gain of 200 is practicable over the audio-frequency range.

Where still higher frequencies have to be handled the gain must once more be limited, as explained in Chapter XV.

## High Frequency Phase Displacement.

Phase shift occurs at high frequency as well as at low frequencies for a similar reason. The presence of $C_o$ across

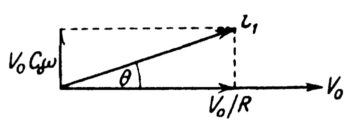



FIG. 59. VECTOR DIAGRAM OF FIG. 58 CIRCUIT AT HIGH FREQUENCIES

$R$ results in a lag, as will be seen from Fig. 59. The current through $R$ is in phase with $V_o$. The total current through $R$ and $C_o$ in parallel, however, leads on $V_o$ as shown and

this is the current supplied by the valve. Hence, $V_0$ leads on $\mu e_g$ by an angle $\tan^{-1} RC_0\omega$, giving a phase shift which increases with $R_1C_0$ and $\omega$.

Once again the effect is cumulative and if the phase shift at the end of an $n$-stage amplifier is not to exceed $\alpha$ degrees, the phase shift per stage, $\theta$, must not exceed $\alpha/n$. This angle $\theta$ is the phase shift in the Miller effect formulae already quoted.

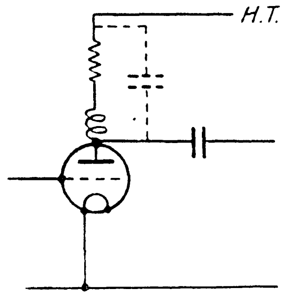



FIG. 60. LOSS OF HIGH FREQUENCIES IN A RESISTANCE-COUPLED AMPLIFIER MAY BE MINIMIZED BY USING AN INDUCTANCE IN SERIES WITH THE RESISTANCE

It is possible to maintain the upper frequencies up to a point by including a choke in series with the resistance as shown in Fig. 60. This maintains a high effective anode impedance at the upper frequencies. Owing to the high resistance the ordinary resonance formulae do not apply. For further information the reader is referred to a paper by G. D. Robinson entitled "Television Receiving Circuits," *Proc. I.R.E.*, June, 1933.

## Direct Coupling.

Where very low frequencies are required, a direct-coupled amplifier may be used. Here the anode of the first valve is

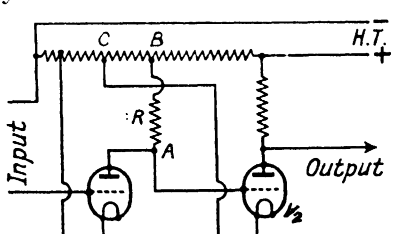



FIG. 61. DIRECT-COUPLED AMPLIFIER

connected direct to the grid of the second as shown in Fig. 61. To avoid the heavy positive grid voltage which

would normally result from such a connection the cathode of the valve is also raised to a positive potential—actually to such a value that the grid is effectively negative to the cathode by the normal grid-bias potential.

The valves must, of course, be of the indirectly-heated type or, alternatively, separate filament batteries must be used. Fig. 61 shows the various connections all tapped off a potentiometer across the h.t. supply. The cathode of $V_2$ is connected to $C$, which is *lower* in potential than $B$. The potential of the anode $A$, however, is also lower than

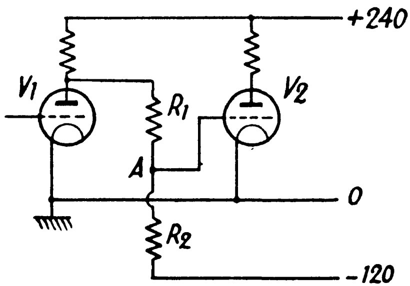



FIG. 62. ALTERNATIVE TYPE OF D.C. AMPLIFIER

$B$ by the amount of the voltage drop on $R$, so that by correct location of $C$ the necessary conditions are fulfilled.

A direct-coupled amplifier will respond to frequencies down to zero (d.c.), but it suffers from the same disadvantages as the ordinary resistance-coupled amplifier at the high frequencies. It is also liable to drift, since small changes of current through the network may upset the somewhat delicate balance of the voltages.

An alternative form of direct coupled amplifier is shown in Fig. 62. Here the voltage developed at the anode of $V_1$ is applied across a pair of resistances $R_1$ and $R_2$, the bottom end of which is connected to a point of negative potential. By suitable choice of $R_1$ and $R_2$, therefore, the potential of the point $A$ may be made slightly negative with respect to the cathode of $V_2$.

The signal voltage is divided in the ratio $R_2/(R_1 + R_2)$ so that there is appreciable loss of gain, but it is usually

easy to obtain more gain from the valve than is actually needed. The method has the advantage that all the cathodes may be at the same potential (and all earthy).

## L.F. Transformers.

Transformer coupling is used in the simpler types of circuit, usually following triode valves. Amplification of

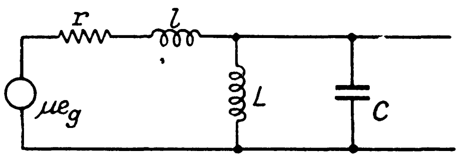



FIG. 63. EQUIVALENT CIRCUIT OF LOW-FREQUENCY TRANSFORMER

the bass frequencies is here dependent on maintaining a high effective primary impedance relative to the valve resistance. This involves using a transformer with a high inductance obtained either by using a large iron circuit (so that the saturation introduced by the heavy anode current shall not be serious) or a high permeability material with a parallel feed circuit. (See Volume I, Chapter XIX.)

The upper frequencies with a transformer are entirely limited by the self-capacitance in the circuit, including the effective grid-cathode capacitance of the succeeding valve. By proper design a resonant effect may again be called into play to assist in maintaining a good response curve.

The ordinary transformer possesses appreciable leakage inductance, the magnetic field of the primary being incompletely linked with the secondary. This may be represented by showing a perfect transformer in series with a small leakage inductance to represent that portion of the magnetic flux which does not link with the secondary.

By ordinary transformer laws the quantities may all be referred to the primary, in which case we can represent the transformer itself by a simple choke.

The equivalent circuit of a transformer, therefore, is as shown in Fig. 63, and it will be seen that there are two

possibilities of resonance. The first of these occurs between the main primary inductance $L$ and the self-capacitance $C$ in the circuit, and is usually arranged to fall between 50 and 100 c/s per sec. in order to maintain good bass response. The second is between the leakage inductance $l$ and the self-capacitance $C$ which occurs at a high frequency, under which conditions the primary inductance acts as an infinite impedance and may be ignored. By suitable design of the leakage inductance this second resonance can be made to occur between 5 000 and 10 000 c/s and thus to maintain the amplification of the upper frequencies.

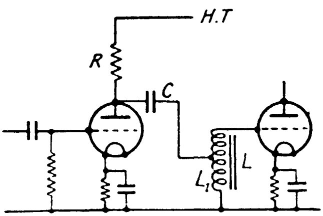



FIG. 64. AUTO-TRANSFORMER ARRANGEMENT

Correct design such as this is an expensive matter and the cheaper makes of transformer are much inferior in their performance, but good amplification with transformer coupling can be obtained up to 10 000 c/s per sec. Auto-transformer couplings are quite frequently used combined with parallel feed circuits and high permeability iron, a typical arrangement being shown in Fig. 64. The capacitance $C$ is arranged to resonate with the tapped portion of $L$ to give resonance around 100 c/s, and from the ordinary laws of resonance it will be clear that the voltage across $L_1$ can exceed the input voltage supplied by the valve, giving a definite rise in the amplification. The dotted curve of Fig. 65 shows the effect without the resonance.

## Cathode Bias.

The circuit of Fig. 64 shows the bias obtained from a resistor in the cathode of the valves. The anode current flowing through this resistor causes a voltage drop which makes the cathode positive to HT— by a small amount. By returning the grid circuit to HT—, therefore, the grid is effectively biased negative by the same amount.

It is important that this bias resistor be properly bypassed. If there is no bypass condenser the a.c. anode

currents due to the signal flowing through the cathode resistor introduce a.c. voltages in the grid circuit in opposition to the applied voltage. This voltage will be $\mu e_g r_c/(r_c + r_a + z)$ where $z$ is the normal anode load, $r_a$ is the a.c. resistance of the valve and $r_c$ is the cathode resistance. If $r_a + z$ happens to be fairly low this negative feed back will be appreciable, and will cause serious loss of gain. It is sometimes deliberately introduced, as explained later, but generally speaking the bias resistor must

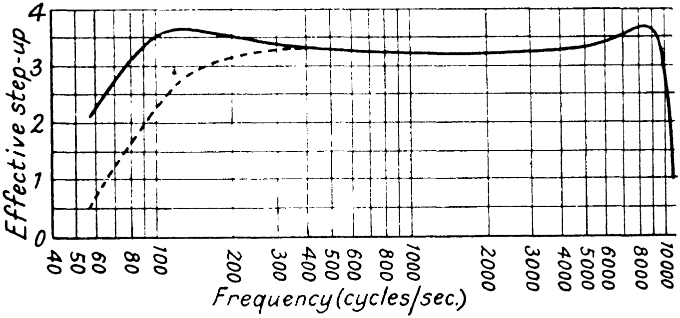



FIG. 65. ILLUSTRATING BASS RESONANCE EFFECT IN AUTO-COUPLED TRANSFORMERS

be by-passed with a condenser which presents a negligible impedance at the lowest frequency to be handled.

Inadequate by-passing is a fruitful cause of loss of the lower frequencies in an amplifier, and if really low frequencies are to be handled the method cannot be used.

## L.F. Compensation.

This loss of gain at low frequencies may be compensated by use of arrangement due to Edwards & Cherry, illustrated in its simplest form in Fig. 66. Here an additional resistance is included in the anode, by-passed with a condenser as with a normal decoupling circuit.

At low frequencies the by-passing of $C_1$ becomes ineffective so that the anode impedance rises and the gain of the stage tends to increase. If these two effects can be made to compensate each other the result should be a uniform stage

gain down to zero frequency. Edwards & Cherry (*Journal I.E.E.*, 1940, vol. 87, p. 178) show that this can be achieved by

(*a*) Making the gain with $C_1$ and $C_2$ omitted equal to the gain with $C_1$ and $C_2$ infinite.

(*b*) Making $Z_1/Z_2 = R\,(g + 1/r_a)$.

This second condition is to ensure a smooth transition between the two states in (*a*), and as this depends upon valve constants the results are somewhat variable. With triodes, also, the values are inconvenient, but the circuit can be made to operate effectively with screen-grid valves provided the screen voltage is held really steady.

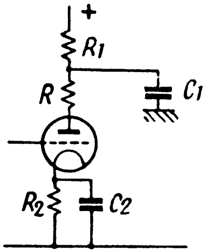



FIG. 66. EDWARDS & CHERRY CIRCUIT

## The Output Stage.

The choice of the valves to be used in the amplifier depends on the voltages to be handled. It is usual to work back from the output stage which is, first of all, designed to work into the optimum load for maximum undistorted output.

The methods of determining this vary with the circumstances and the degree of accuracy required. The most general method is to adopt the graphical tactics outlined in Volume I, Chapter XIX, and it will be as well to review the process briefly.

Fig. 67 shows typical characteristics for a triode. The working point $O$ is chosen in this case corresponding to 200 volts on the anode and $12\frac{1}{2}$ volts grid bias. The variations of anode voltage and current may be determined by drawing a line through the point $O$ having a slope corresponding to the effective resistance of the load, and for absolutely distortionless working the points where this *load line* cuts the limiting characteristics should correspond to equal excursions of anode voltage and current from the mean value.

In Fig. 67, for example, the mean bias is $12\frac{1}{2}$ volts, so that the limits of grid swing will be 0 and 25, assuming the valve to be fully loaded, and the distances $OA$ and $OA'$ should be equal if the load represented by $AA'$ is correct.

Actually they are far from equal, showing that considerable distortion is occurring. This particular distortion, wherein one half of the wave is flattened and the other peaked, is called *second harmonic distortion* because the form of wave is the same as would result if a pure wave were mixed with a (smaller) wave of twice the frequency as shown in Fig. 68.

It will be clear from Fig. 67 that a non-linear operation

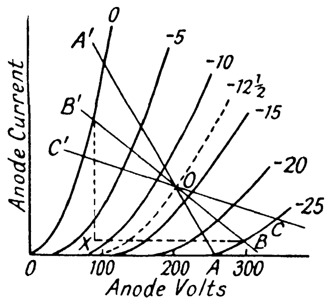



FIG. 67. ILLUSTRATING CHOICE OF OPTIMUM LOAD FOR A TRIODE

The figures on the curves refer to the grid voltage

such as this will be accompanied by an increase in the mean anode current (because the current swing $OA'$ is greater than $OA$). To allow for this the second harmonic in Fig. 68 has been shown displaced above the zero line for the fundamental.

Now let $V$ be the maximum amplitude of the fundamental and $dV$ that of the harmonic. It will be seen that the positive peak of the combined wave is $V + 2dV$ and that of the negative wave is $V - 2dV$. Hence the ratio of the positive to negative peaks is $(V + 2dV)/(V - 2dV)$, i.e. $OA'/OA$ in Fig. 67. If $d = 0{\cdot}05$ (5 per cent distortion) this is $1{\cdot}1/0{\cdot}9 = 1{\cdot}22$. Hence if $OA'/OA$ is not greater than $1{\cdot}22$ we shall not introduce more than 5 per cent distortion and this criterion is often used by circuit designers.

The load line $A'OA$ does not fulfil this requirement, but $B'OB$ does. It will be noted that the anode current is no longer reduced to zero at each swing, so that the valve is

operating less efficiently, while the power output (which is $\frac{1}{8}XB \cdot XB'$ as explained in Volume I) is also less than before. As we reduce the slope of the load line the power output falls off still further while the distortion at first becomes less and then begins to increase again, due to the tendency of the characteristics to become flatter with high grid bias.

A steep load line corresponds to a low load resistance. Hence, we find with a triode that as we increase the load

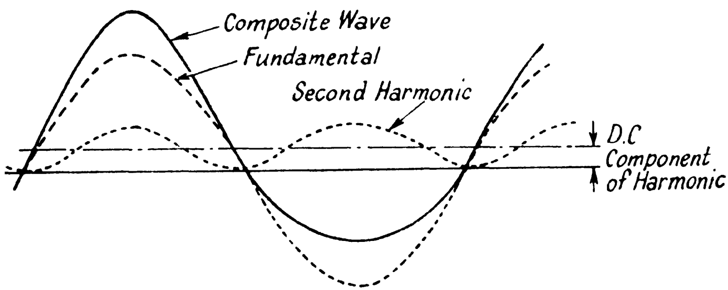



FIG. 68. ILLUSTRATING FLATTENING OF WAVE DUE TO SECOND HARMONIC

the power output rises to a maximum and then falls off again, while the distortion falls to a minimum and then begins to rise again. The optimum distortion condition does not coincide with that for maximum power output.

## Output Tetrodes and Pentodes.

A similar though slightly modified procedure is adopted with pentodes, and the critical-distance or beam tetrodes which are superseding them. The latter valves have the anode located farther from the screen and so disposed that the usual secondary-emission kink is removed. At the same time the screen current is appreciably reduced by aligning the screen and control grid so that a more efficient valve results.

In use, the valves behave similarly to a pentode, and a typical set of characteristics is shown in Fig. 69. The line $AOA'$ represents a condition giving no second harmonic distortion at all, for $AO = OA'$. This, however, is not the best condition, for there is a marked third harmonic com-

ponent which can be detected by noting the increments in anode current for equal increments of grid voltage. These are shown in the table below, which also shows the change in current per 5 volts grid change.

TABLE OF ANODE CURRENTS IN FIG. 69

| Grid Bias | Load Line $AA'$ $I_a$ | Load Line $AA'$ Increments | Load Line $B'B$ $I_a$ | Load Line $B'B$ Increments |
|---|---|---|---|---|
| − 30 | 55 | | 52 | |
| | | 15 | | 18 |
| − 25 | 70 | | 70 | |
| | | 35 | | 30 |
| − 20 | 105 | | 100 | |
| | | 30 | | 35 |
| − 15 | 135 | | 135 | |
| | | 30 | | 35 |
| − 10 | 165 | | 170 | |
| | | 35 | | 35 |
| − 5 | 200 | | 205 | |
| | | 15 | | 35 |
| 0 | 215 | | 240 | |

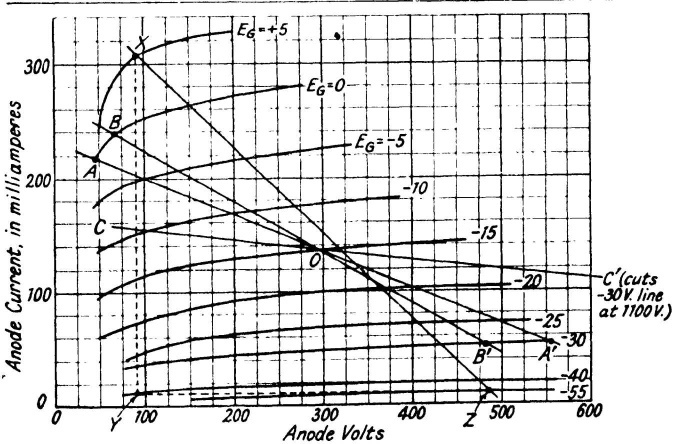



FIG 69. CHARACTERISTICS OF A TYPICAL OUTPUT TETRODE

It will be seen then working from the middle (— 15 V.) we have, on each side, changes of 30, 35 and 15 mA. respectively for successive increments of 5 volts. This will give a wave with a marked third harmonic, as shown in Fig. 70.

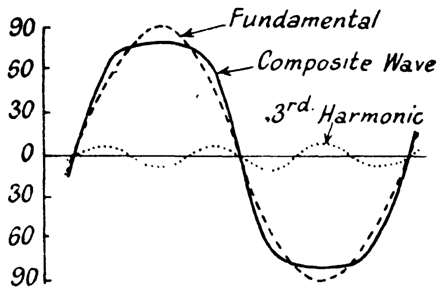



FIG. 70. ILLUSTRATING INFLUENCE OF THIRD HARMONIC ON WAVEFORM

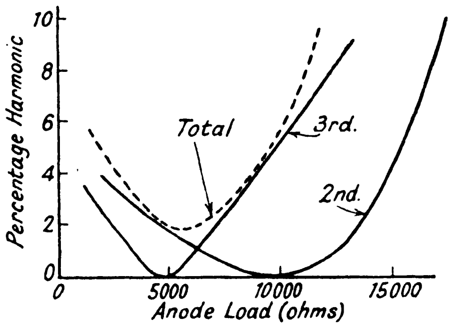



FIG. 71. ILLUSTRATING THE EFFECT OF HARMONICS AND THE OPTIMUM LOAD POINT

Note that with third harmonic the wave is symmetrical. This is true of any odd harmonic. On the other hand, if we reduce the load slightly to that shown by $BOB'$ we introduce about 5 per cent second harmonic for $OB = 1{\cdot}26OB'$, but over most of the line equal increments of grid voltage give equal anode current changes, so that we have no third harmonic.

This behaviour is characteristic of pentodes and tetrodes. As we increase the load the second harmonic content falls to zero and then rises again (this time due to a flattening of the top of the curve instead of the bottom). The third harmonic behaves similarly but falls to zero much earlier and we usually find that at the point of zero second harmonic the third has risen alarmingly. The best condition is between the two as shown in Fig. 71.

It is difficult to locate the best load from the characteristics but curves similar to Fig. 71 are usually available from the makers.

As an approximation the following procedure may be used. First determine

$\alpha$ = current swing for a positive grid swing = 0·71 times the peak grid swing.

$\beta$ = current swing for a negative grid swing of the same value.

Then evaluate three parameters as under

$$A = \frac{\text{positive peak current}}{\text{negative peak current}}$$

$$B = 1 + \alpha/\beta.$$

$$C = 1{\cdot}41\beta/\text{negative peak current}.$$

Then

$$\%\ \text{2nd harmonic} = \frac{A - 1}{1 + A + BC} \times \frac{1}{100}$$

$$\%\ \text{3rd harmonic} = \frac{1 + A - BC}{1 + A + BC} \times \frac{1}{100}$$

## Impedance Limiting.

It should be noted that tetrodes or pentodes should not be run with high loads. The line $COC'$ represents such a load, and it will be seen that if the full $\pm$ 15 volts grid swing is applied the anode voltage will swing to a voltage of several thousand, which will certainly cause a breakdown somewhere.

A safety load of several times the normal load should therefore be connected permanently across the output or else a safety spark gap incorporated.

The average loud speaker impedance, while fairly constant over the middle registers, rises rapidly in the upper frequencies. This causes the load line to swing round, giving rise to the same trouble if a large high (audio) frequency input is applied. Even if no damage results, very bad distortion will be produced. To avoid this a limiting load is often connected in parallel sometimes with a condenser in series, the condenser value being chosen such that its reactance is high at normal frequencies, but falls rapidly above a few thousand cycles thereby bringing the limiting resistance into play.

## Choice of Valves.

Having decided upon the valve to obtain the required output, one knows the grid input required. If sufficient voltage can be obtained direct from the detector this is all that is necessary. Valves are made to-day which will give two to three watts out with four volts peak signal input, and this value can easily be obtained from a diode detector. Some commercial receivers use this system.

In general, however, one l.f. amplifying stage is desirable. This may be the detector itself if an amplifying detector is used. If not, a first l.f. valve must be employed capable of providing this voltage. Due to the coupling impedance in the anode circuit the anode voltage will, in general, be less than the full amount, particularly if resistance coupling is used, and the valve must be capable, with the actual effective anode voltage, of providing without distortion an anode swing equal to the voltage required. This involves as a corollary that to provide this anode swing the grid swing required on the input to the valve must be within the limits of the grid bias used.

It must be remembered that, in a design of this type, allowance has to be made for the peak condition, i.e. the strongest possible signal likely to be handled under full modulation.

## Class B Operation.

The ordinary low-frequency amplifier operates as a

Class A amplifier, being biased approximately to the middle of the straight portion of the characteristic. This involves a large steady anode current which is the same whether the signal is large or small. Particularly with battery-operated receivers, this involves a waste of current, since the periods when anything like full modulation is in operation are only a small fraction of the total time.

This has led to the introduction of Class B amplifiers for audio-frequency purposes. As explained in the chapter on transmitters, this form of operation consists in biasing the valves to such a point that the anode current is nearly zero. Positive half-cycles of grid voltage then cause the anode current to swing over the straight portion of the characteristic, while the negative half-cycle produces no appreciable effect.

In a transmitter the anode voltage maintains itself over this portion of the cycle by virtue of the tuned circuit in the anode circuit. With a low-frequency amplifier this is not permissible because we are dealing with a variety of frequencies and any undue resonance cannot be tolerated. It is necessary, therefore, to use two valves in push-pull, one of which handles one half-cycle and the other the next. The outputs are combined in the normal manner. A typical circuit is shown in Fig. 72.

For battery operation two pentodes may be used, each biased to about twice the normal value and swinging on each alternate half cycle from this value up to zero bias. Special Q.P.P. (quiescent push-pull) valves are made for this purpose, having two pentodes in one envelope. The two elements are balanced to take approximately equal shares of the drive by individual adjustment of the screen volts.

Class B valves are also used occasionally. These are triodes designed to run into the positive grid region in which case their characteristics become similar to those of a pentode. Such a combination absorbs power on its input and has therefore to be fed with a small power valve operating through a *driver transformer* which is designed in the same way as an output transformer to feed into the load provided by the grid current drive.

Class B triodes, however, cannot conveniently be matched except by selection and the system is not greatly used. It is employed, however, with pentodes or tetrodes not only for battery operation but also with mains valves. The arrangement is similar to the Q.P.P. scheme mentioned above and may, indeed, be used in this way without running into grid current. This is sometimes called class AB1 working. Greater output is obtained, however, by running into grid current at the peaks of the swing, the

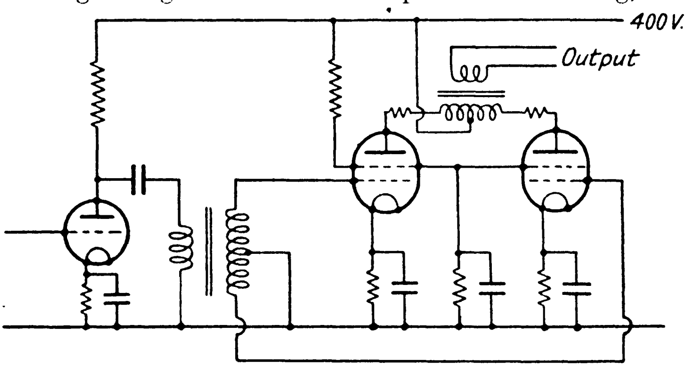



FIG. 72. TWO TETRODES ARRANGED FOR CLASS AB2 OPERATION (POSITIVE DRIVE)

preceding valve being designed to supply the power necessary.

Fig. 72 shows two tetrodes arranged to operate in this Class AB2 (positive drive) condition. The output of the two valves in normal Class A push-pull is 14·5 watts with 9 per cent total distortion. Under Class AB2 conditions 60 watts output may be obtained with only 3 per cent distortion. It is important to keep the screen voltage constant, and a separate power unit is often used for the purpose.

The operating condition of each valve is shown at *XZ* in Fig. 69. The valve normally operates at 25 volts bias and runs up to + 5 volts. As the valve runs towards zero its pair runs negative. The total swing, therefore, is from − 25 to + 5 on one valve and from − 25 to − 55 on the

other. The peak power output is thus represented by the triangle $XYZ$. $YX$ is the peak current swing and $YZ$ the peak voltage swing so that the r.m.s. power is $XYZ/2$. This is the total power for both valves and will be seen to be 59·5 watts.

In calculating distortion, the anode current change must be taken as the sum of the changes in each valve. E.g. a 5-volt increase from 25 to 20 will give $100 - 68 = 32$mA, while the other valve of the pair will contribute $68-52=16$, making a total of 48mA. The change from $E_g = 0$ to $E_g = +5$ gives $310 - 262 = 48$mA, which is the full change since in this condition the second valve is practically cut off. Similar calculations at intermediate points give similar current changes showing that with the conditions chosen the operation is practically distortionless.

## Feed-back.

An important factor in amplifier design and operation is that of feed-back. Components or even leads at the output end of the chain must not be allowed near the input. It is also desirable to filter out any high-frequency currents immediately following the detector. This must be done as efficiently as possible, as otherwise amplified h.f. voltages will appear in the output stage and will radiate from the loudspeaker back on to the earlier portions of the set, causing instability. A large h.f. component in the audio-frequency output will also cause the valves to overload, even though they are quite capable of handling the normal low-frequency signal.

Finally, attention must be paid to the question of common-impedance coupling already referred to in connection with h.f. amplifiers. These decoupling circuits must be used as explained in Volume I, and if, as it is usual, resistances are used for the decoupling, due allowance must be made for the voltage drop which will occur on these resistances. The valves must still be capable of handling the full anode swing with the reduced anode voltage.

Decoupling is also used in grid circuits at times, and also a.v.c. circuits, in both of which cases the d.c. potentials only are of consequence. The voltage is supplied through

a high resistance of the order of 1 megohm, bypassed at the far end with a large condenser.

Miller effect is less important with the transformer type of amplifier. The anode circuit is inductive, except around the two main resonance joints. Hence, the input capacitance of the valve preceding it is not unduly high (except at the resonant points) while the input resistance tends to be negative, i.e. a certain reaction effect is obtained. While this does not usually cause self-oscillation, except in amplifiers having more than one stage, it does adversely affect the response curve and must be borne in mind.

A loud-speaker acts mainly as a resistive load, since it is absorbing power, and the output valve reflects an appreciable capacitance across the circuit preceding it unless a screened valve is used. This is one of the chief arguments in favour of pentode or tetrode output valves where, owing to the greatly reduced internal capacitance, the feed-back is rendered of small proportions only.

**Negative Feed-back.**

A development of importance has been the use of negative feed-back. This is a feed-back in a direction which reduces the amplification and at first sight would appear to have little advantage. If the feed-back is sufficient, however, a surprising effect is obtained.

Consider an amplifier with a gain $A$. Let the input be $v$. Then the normal output is $vA = V$. Now transfer a portion of this output $\beta$ back to the input. The gain of the amplifier will be reduced and to obtain the same output as before we must increase the input by an amount $\beta V$. The effective input voltage thus becomes $v + \beta V = \frac{V}{A} + \beta V$

$$= V\left(\frac{1}{A} + \beta\right).$$

The effective gain of the amplifier is thus $V/V\left(\frac{1}{A} + \beta\right)$

$= \frac{A}{1 + \beta A}$ which can be rewritten $\frac{1}{\beta} \cdot \frac{1}{1 + \frac{1}{\beta A}}$

Now if $\beta A$ is large, $1/\beta A$ tends to zero, so that the gain tends to become simply $1/\beta$.

This means that the gain of the amplifier is constant and is *independent of the components of the amplifier, including the valves.* If 1/100th of the output voltage is fed back, the gain is 100, provided that $\beta A$ is large compared with unity, which means that in the case chosen $A$, the normal amplifier gain, must be of the order of 1 000.

We have, of course, obtained this independence of valves

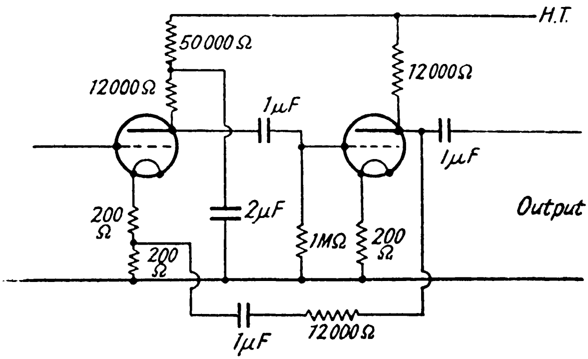



FIG. 73. NEGATIVE FEED-BACK CIRCUIT

and circuit at the expense of a good deal of the gain—9/10ths to be exact, but the advantages may outweigh this fact. Experiments made by the author* have shown that the results bear out theory in this respect and that deliberate distortion can be offset practically completely, and the same output obtained provided the input is increased to allow for the reduced gain.

In practice, negative feed-back is usually used to a partial extent only, the circuit being made tolerably good and then improved by feed-back. Fig. 73 shows a negative feed-back circuit. In each valve a small feed-back is obtained by omitting the customary by-pass condenser

* "Linearity and Negative Feed-back," *Wireless World*, 4th August, 1938.

across the bias resistor, while in addition there is a small deliberate feed-back from output to input. Various other forms of circuit can be devised, and it is easy to decide whether the feed-back is in the right direction.

The action breaks down if the phase of the feed-back suffers any serious shift in its transference from back to front, or if the amplifier itself has any serious phase shift. It is, in fact, possible at high frequencies for the total phase shift to be more than 90°, in which case the feed-back becomes positive and self-oscillation will result. This possibility must always be guarded against.

**Effect of Feed-back on Output Impedance.**

The feed-back may be proportional to the output voltage as in Fig. 73 or to the output current. This could be done by introducing a suitable small resistance in the earthy side of the output circuit itself. The voltage developed across this resistance would then be proportional to the output current.

Voltage feed-back tends to maintain the output voltage constant. Now a generator develops at its terminals a voltage equal to the internal e.m.f. less the voltage drop on the internal impedance. Hence if the voltage output is to be constant the internal impedance must be zero, and the more we reduce the internal impedance the closer we shall approach this condition. Voltage feed-back, therefore, has the effect of reducing the effective internal impedance of the output stage, which is often convenient.

A pentode, for example, has a high internal impedance, tending in fact to be a constant current device which is a condition associated with infinite internal impedance. Suitable voltage feed-back will convert the circuit to one having characteristics approaching constant voltage output, with a consequent marked reduction in internal impedance. The circuit in fact will have triode characteristics.

Conversely, current feed-back will tend to a constant current condition and can be arranged to make a triode exhibit the high internal impedance of a pentode. By a mixture of the two the internal impedance of the output stage can be adjusted to any specified value.

## Cathode-follower Circuit.

A particular form of negative feed-back circuit now in common use is the cathode follower. The practice of including a resistance in the cathode lead of a valve, in order to develop a suitable grid-bias voltage, is well known. In this usage, however, the cathode resistor is bypassed with a condenser so that the a.c. component of the anode current shall not develop any appreciable voltage in the cathode circuit.

If this bypass condenser is omitted as in Fig. 74 (*a*), the

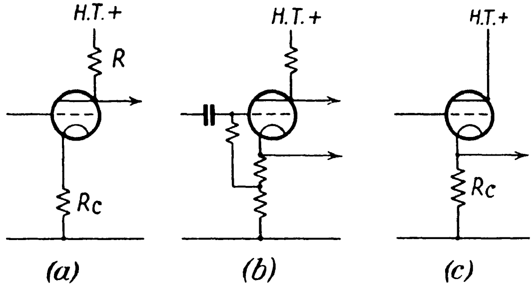



FIG. 74. DEVELOPMENT OF THE CATHODE-FOLLOWER CIRCUIT

circuit becomes a negative feed-back system. If $i_a$ is the a.c. component of the anode current, we can write

$$i_a = -\mu e_g/[r_a + (R + R_c)]:$$

$$V_R = i_a R;$$

$$V_{in} = e_g - i_a R_c.$$

$$\text{Stage gain} = i_a R/(e_g - i_a R_c) = R/[(e_g/i_a) - R_c].$$

Substituting the expression above for $i_a$, this reduces to

$$A = \mu R/[r_a + R + R_c(1 + \mu)].$$

This is the same as the normal expression except that $R$ in the denominator is increased by $R_c(1 + \mu)$, with consequent reduction in gain (and, possibly, improvement in linearity).

The cathode resistor may be increased beyond the value required for bias, the grid leak being tapped across a

portion only, as in Fig. 74 (*b*). Under such conditions, apart from the reduction in gain due to feed-back, there is a further loss of output because the anode resistance is only a fraction of the total external load, and an appreciable fraction of the output voltage appears across the cathode resistor. This effect is sometimes utilized to obtain two output voltages in opposition—one from the anode and the other from the cathode.

It will be clear that an increased anode current will cause an increased voltage drop on both anode and cathode resistors. But since the anode is already negative with respect to earth (HT+) it will become still more negative. The cathode is positive to earth and will increase its positive potential, so that anode and cathode potentials move in opposite directions.

The voltage across the cathode resistor is $i_a R_c$ so that the stage gain becomes

$$A = \mu R_c / [r_a + R + R_c(1 + \mu)].$$

If we reduce the anode resistance to zero, as in Fig. 74 (*c*), this reduces to

$$A = \mu R_c / [r_a + R_c(1 + \mu)]$$

and in the limit, if $R_c(1 + \mu)$ is much greater than $r_a$ we have simply $A = \mu/(\mu + 1)$.

Such a circuit is called a cathode follower. It will be seen that the maximum gain is just short of unity and may not be more than a fraction if $R_c(1 + \mu)$ is of the same order as $r_a$. The arrangement, in fact, operates not as an amplifier but as an impedance changer, which is often of great convenience.

For example, a high-resistance potentiometer cannot be used to feed the grid of a valve if the frequency is high because the parallel reactance of the valve input and stray capacitance will be small. This largely short-circuits the bottom portion of the potentiometer so that it ceases to provide adequate control. If, however, the input is applied to the grid of a cathode follower, the cathode resistance may be in the form of a potentiometer of quite low resistance such that the capacitance of the succeeding stage does not

affect the operation. The input impedance to the cathode follower, on the other hand, is very high being simply the grid-cathode capacitance (including strays). Miller effect is not present.

The circuit is also often used as a buffer stage to prevent one circuit from interacting with another.

## Feed-back Networks as Tuned Circuits.

An important development of feed-back technique is the use of the system to provide the characteristics of a tuned circuit. With this arrangement the feed-back is routed

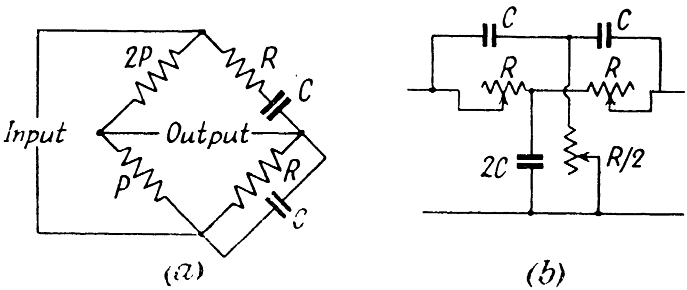



FIG. 75. FREQUENCY BRIDGE NETWORKS

through a bridge of which the balance is dependent on frequency. Normally the bridge will be unbalanced, so that there will be a considerable fraction of the original feed-back voltage available at the output terminals of the bridge. This voltage will be applied to the amplifier and will reduce its gain to some small value.

At the particular frequency for which the bridge balances, however, the feed-back will become very small, thus releasing the amplifier and permitting it to develop nearly its full gain. The response of the amplifier may thus increase a hundredfold at this frequency of "tune," while for a short range on either side there will be an intermediate response, the whole result resembling the resonance curve obtained with a normal tuned circuit.

One advantage of the system is its high stability, condensers and resistances having a lower temperature coefficient than inductances while, even more important, the

frequency is not affected by changes in the amplifier but is solely dependent on the bridge network in the feed-back circuit. It is also very convenient for low frequencies where the inductances required with normal methods become very bulky, while a third and most useful feature is that

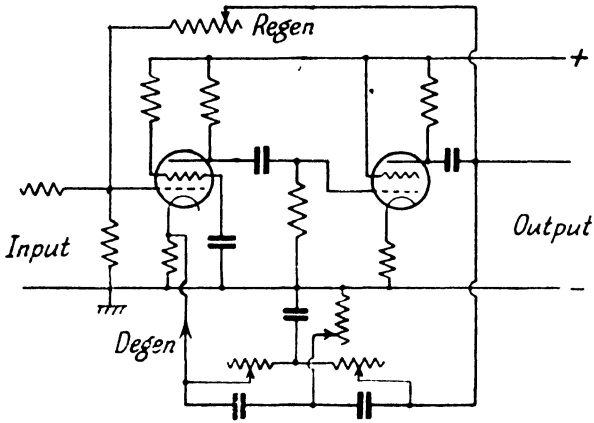



FIG. 76. FEED-BACK OSCILLATOR

the frequency is very easily made variable by altering the bridge network.

Fig. 75 (*a*) shows a simple frequency bridge which balances when $f = 1/2\pi CR$. In this form it is inconvenient, however, for both the output terminals are at different potential to the input. A star-mesh transformation (see Appendix), however, enables the circuit to be re-arranged in the form of Fig. 75 (*b*).

Here one side of both input and output is earthy. Fig. 76 shows a simple selective amplifier of this type using the 3-terminal bridge of Fig. 75 (*b*).

## Regeneration—Feed-back Oscillators.

The sharpness of tune is only fair, a typical network being equivalent to a tuned circuit having a $Q$ of 25. This is quite good for low frequencies but not as good as can be obtained with good tuned circuits. The application

of a little *positive* feed-back, however, enables the tuning to be sharpened considerably and a $Q$ of 700 is easily obtainable. Such a system, however, requires careful handling and suffers from the usual defects of a reacting system, the most important being that the system is no longer independent of the amplifier, both amplification and, to a lesser extent, frequency being influenced by any changes in the amplifier or its supply voltages.

If the positive feed-back is increased beyond a certain point the circuit will oscillate continuously at the frequency to which the bridge is set. This, therefore, provides a convenient form of oscillator which may be made to have a high degree of stability and cover a wide range of frequency and such oscillators are becoming increasingly popular. Frequencies from a few c/s up to several kc/s can be handled successfully.

The important feature in the design of feed-back systems as a whole is that there shall be negligible phase shift in the amplifier *within the operating range of frequencies*. Outside this range phase shift may be permitted, but it should not be allowed to exceed $\pi/4$ until the gain of the amplifier has fallen below unity. Loss of gain and phase shift go together (see p. 114) and it can easily be demonstrated that the less the number of couplings the more easily can this condition be complied with. In fact with one stage only the phase shift can never exceed $\pi/4$, so that it is preferable to arrange the selective feed-back around one stage only, adding further stages quite independently if more gain is necessary.

## Phase-shift Oscillators.

Reference may be made to another form of oscillator which is occasionally used. Here three stages are used, so designed that a phase shift of $2\pi/3$ per stage is obtained. The output is then coupled back to the input causing continuous oscillation.

This will only occur at the particular frequency for which the total phase shift is $2\pi$, so that the voltage feed-back is in phase with the input. If the constants of the interstage couplings are altered the frequency of the oscillation automatically changes until this condition is again complied with.

Hence once again we have an oscillator using resistances and condensers only of which the frequency is variable over a wide range by varying the values of resistance and/or condenser. Both types of oscillator generate a very pure waveform if precautions are taken to avoid overloading of the valves.

## CHAPTER VIII

## THE RECEIVING AERIAL

THE requirements for a receiving aerial are slightly different from those at the transmitting end. If the object were merely the production of the most efficient energy-collecting system, the receiving aerial could be a duplicate of that of the transmitter (without the necessity for high insulation). The problem in receiving, however, is to tune in the required signal, which is usually weak, and to disregard much stronger signals from relatively nearer transmitters operating on wavelengths which may not be much different from that of the desired station.

Now an aerial operating at its natural wavelength is very broadly tuned. Moreover, receiving aerials are usually required to operate over a range of wavelengths, except in certain special cases. For ordinary medium- and long-wave reception, therefore, the receiving aerial is only a fraction of a wavelength long. It is made high so that the voltage induced, which is the product of the field strength in microvolts per metre, and the effective height in metres, shall be as large as possible.

### Aerial Coupling Circuits.

The receiving aerial is coupled to the receiver through a transformer, usually of the inductive- or direct-coupled type as shown in Fig. 77A. The function of this arrangement is twofold. In the first place, it increases the voltage applied to the first valve of the receiver by means of the step-up in the transformer, and at the same time it improves the selectivity, because the aerial is possessed of a relatively large resistance, mainly radiation resistance, and since the aerial-tuned circuit combination is merely a variety of coupled circuit any resistance in the aerial itself is reflected into the tuned secondary.

A further point is that the aerial capacitance, which may be anything from 0·0001 to 0·001 $\mu$F. for the average receiving aerial, is effectively connected across the tuning circuit.

The secondary circuit is usually tuned with a condenser of about 0·0005 $\mu$F., and if there is a permanent capacitance of a similar value already across the circuit, the capacitance range is only about 2 : 1, giving less than $1\frac{1}{2}$ : 1

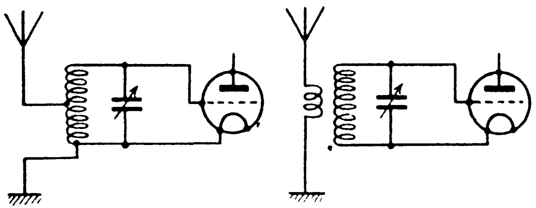


FIG. 77A. TWO TYPES OF INDUCTIVE AERIAL COUPLING

variation in the wavelength. If the aerial is transformer-coupled, the effective capacitance across the full circuit is reduced and, consequently, the tuning range is not seriously restricted.

The equivalent circuit of a receiving aerial circuit is shown in Fig. 77B, and the approximate treatment is as

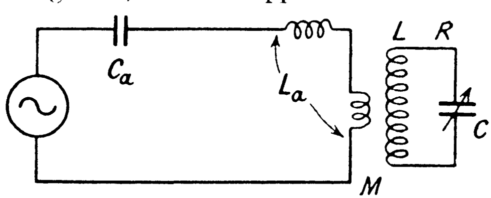



FIG. 77B. EQUIVALENT AERIAL CIRCUIT

follows. Aerial resistance is neglected in comparison with the equivalent resistance reflected into the circuit from the secondary, since this is usually high.

Where the secondary is tuned, the equivalent resistance in the aerial circuit is $M^2\omega^2/R$.

Hence, $$i_1 = \frac{e}{M^2\omega^2/R + j(\omega L_a - 1/\omega C_a)}$$

where $L_a$ includes the primary inductance of the transformer;

$e_2 = M\omega i_1$ and $E$, the voltage across the secondary coil, $= M\omega \, . \, (L\omega/R)i_1 = (M/RC)\ i_1$ since $\omega^2 = 1/LC$.

Hence the effective step-up is

$$\frac{E}{e} = \frac{M/RC}{M^2\omega^2/R + j(\omega L_a - 1/\omega C_a)}$$

$$= \frac{M/C}{M^2\omega^2 + j\,X_1 R}$$

where $X_1$ is the reactance of the aerial $= \omega L_a - 1/\omega C_a$.

Now the aerial circuit is not tuned. If it were, double-

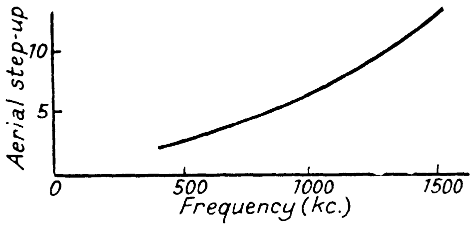



FIG. 78. TYPICAL AERIAL STEP-UP CURVE

hump effects would occur and the tuning would be upset. Hence, the tune of the aerial circuit is always kept well outside the tuning range, and the optimum coupling is obtained when the two terms in the denominator are equal—i.e. $M^2\omega^2 = X_1 R$.

This usually gives a step-up of the order of five or ten to one. It varies with the frequency, being higher with increased frequency. A typical aerial coupling step-up curve is shown in Fig. 78.

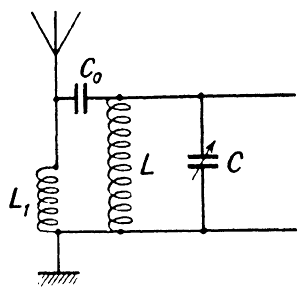



FIG. 79. CIRCUIT FOR OBTAINING UNIFORM AERIAL STEP-UP

A capacitance coupling would give the reverse effect, and aerial circuits are sometimes arranged to combine both forms of coupling in order to obtain a uniform transfer of energy over a range of frequency. Fig. 79 shows such a circuit. The inductance of $L_1$ is higher than usual, so that it would tune with $C_a$ just above the highest wavelength to be received.

It is loosely coupled to the secondary and the energy is transferred partly by inductive coupling and partly directly to the top of the secondary through $C_0$.

The direction of $L_1$ is such that these two couplings are in opposition. At low frequencies the energy transfer is nearly all inductive since $C_0$ is small and offers a large impedance to the currents. As the frequency rises, more voltage is induced in $L$ from $L_1$, but more current flows through $C_0$ in opposition and so keeps the total secondary voltage fairly constant.

Various other special forms of aerial coupling are used, but it is not possible to discuss them further here.

**Band-pass Coupling.**

In order to obtain increased selectivity the aerial and the secondary are sometimes separately tuned. When this is done the system becomes a species of coupled circuit, and the coupling must be kept below a certain limiting value for satisfactory operation.

The critical coupling is obtained when the effective resistance introduced into the primary by the secondary circuit coupled to it is equal to the initial primary resistance, which occurs when $M\omega = \sqrt{(R_1R_2)}$ (see Chapter IX). Under these conditions the maximum current is obtained in the secondary and the maximum energy transferred from the primary.

With critical coupling the secondary resonance curve shows only one peak, while that of the primary shows two peaks very close together. As the coupling is increased the primary peaks move farther apart, and a double hump also appears in the secondary resonance curve. The frequencies of the peaks are discussed in the next chapter.

Receiving circuits are sometimes operated in this condition, the peaks being 5 to 10 kc/s apart so that the side-bands are received at a strength comparable with the carrier, whereas with the ordinary peaked resonance curve the frequencies 5 to 10 kc/s off tune are appreciably attenuated, particularly with a sharply tuned circuit. This gives rise to a loss of the upper frequencies, or "top cut," which is minimized by band-pass tuning. Beyond

the limit of the peaks the current falls very sharply, giving a better selectivity than could be obtained with one circuit alone.

For broadcast reception it is usual to tune the two circuits with a twin condenser. To obtain a reasonable similarity between the characteristics the aerial is coupled to the primary, just as with a simple coil. The two sections

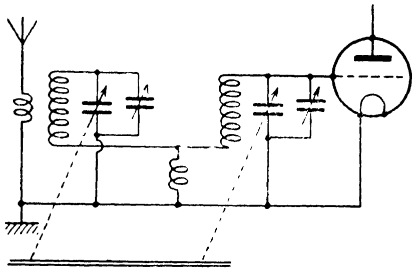


FIG. 80. BAND-PASS AERIAL CIRCUIT

of the condenser are identical, but small semi-fixed capacitances, known as *trimmers*, are mounted on the side and are adjusted to compensate for the differing self-capacitances (including the effect of the aerial). Fig. 80 shows a simple inductively-coupled band-pass filter.

The displacement of the peaks depends on the frequency as explained on page 160. With a magnetically-coupled system the peak separation increases with the frequency. It is possible to couple the circuits capacitatively, in which case the reverse action occurs. "Mixed" filters are often used to combine both effects and obtain a constant band width.

For further treatment of this subject the reader is referred to *Wireless Engineer*, Vol. IX, pages 546 and 608, October and November, 1932.

## Directional Aerials.

For point-to-point working directional aerials are often used. The extent of the directivity depends upon a variety of circumstances. Directional aerials for medium and long waves are usually rather difficult to arrange effectively, whereas on short waves the problem is comparatively easy.

The principal types of directional aerial are discussed in Volume I, Chapters XXII and XXIII, the method usually adopted being to use an arrangement of spaced aerials so that there is an appreciable time lag between the voltages induced by the waves at the two ends of the system. By suitable phasing of the currents set up in the aerial this time lag can be caused to accept a signal coming from one direction and to reject one coming from the opposite direction.

**The Beverage Aerial.**

A form of directional aerial which is particularly useful on short waves is that developed by C. W. Beverage in America some years ago. It consists of a long horizontal

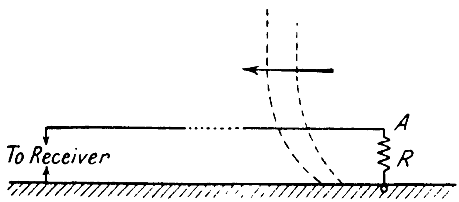



FIG. 81. SHOWING HOW WAVE DRAG IS UTILIZED IN THE BEVERAGE AERIAL

wire a few feet off the ground pointing in the approximate direction from which the signal is coming. It operates by virtue of the fact that the lower portion of the vertical field in a wireless wave drags slightly due to the resistance of the earth, as shown in Fig. 81.

There is thus an horizontal component which induces voltage in the wire. The voltage in the first element will travel along the wire to the far end. Meanwhile the advancing wave is inducing further voltages in the successive elements of the wire and it is clear that these will be roughly in phase. The time taken by the wave to travel down the wire is longer than that taken by the wave in free air, but the effect is nevertheless cumulative and if the wire is several wavelengths long a considerable signal builds up at the far end.

Signals from the reverse direction produce the same result but on arrival at the point $A$ they are absorbed by the resistance $R$ which is made equal to the characteristic impedance of the line and thus absorbs without reflection. Signals from any other direction clearly cannot build up.

**Diversity Reception.**

Reference may be made to a particular form of receiving aerial which is intended to overcome fading. As explained

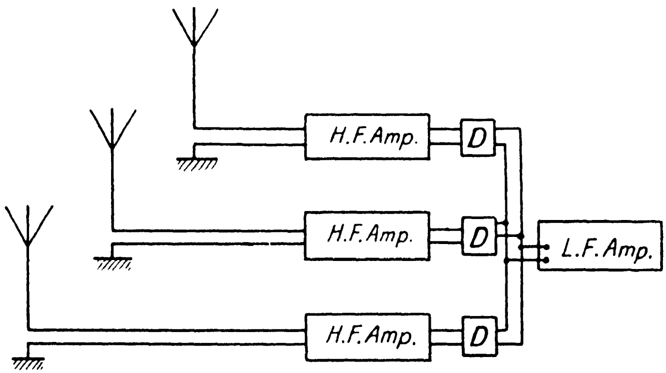



FIG. 82. ARRANGEMENTS FOR DIVERSITY RECEPTION

in Volume I, fading is produced by interference with the travel of the waves during their reflection at the Heaviside Layer. In particular, the plane of polarization is twisted so that a wave may arrive horizontally polarized, in which case it will have no effect on the ordinary vertical receiving aerial.

This variation in the plane of polarization may be spasmodic or it may be regular, arising in the latter case from a circularly polarized wave in which the plane of polarization is rotating continuously.

Attempts may be made to minimize fading by the use of automatic volume control, but this is not entirely satisfactory, for no amount of increased amplification can compensate for a complete fade-out, while, in addition, the increasing amplification always brings with it an increased background noise.

It is found, however, that the reception at different localities is not the same at the same instant. When the signal has faded to vanishing point in one locality it may be quite strong only a short distance away.

This is understandable in view of the explanation just given of the cause of fading, for if we consider a circularly polarized wave its plane of polarization at any instant will depend upon the distance of the receiving point from the Heaviside Layer.

Consequently, if two or more receiving aerials are erected a few wavelengths apart and each one individually tuned to the required signal, and the output of each mixed subsequent to rectification, we should obtain a reasonably uniform signal. When one aerial is receiving practically nothing there will be some signal in one at least of the others, and in practice such a combination usually provides satisfactory reception.

Certain difficulties are experienced in applying this system to telephone reception, because the modulation is not always in phase on the three receivers. To overcome this, square-law detectors are used, together with an automatic volume-control device following the detectors, which operates on all three receivers simultaneously. Since the square-law detector operates much more effectively on a strong signal than on a weak one, this arrangement ensures that the greater part of the output obtained from the system comes from the particular aerial which is receiving the best signals at that moment. The system is reasonably satisfactory in practice and is used to a considerable extent.

The aerials are spaced a few wavelengths apart and the signals are brought therefrom by means of radio-frequency feeder lines of the type discussed in Chapter III. The aerials themselves may be of any suitable type, Beverage aerials being used in many instances.

## Frequency Diversity.

A somewhat simpler form of diversity reception, sometimes used for telegraphic signals, depends upon the fact that fading depends to a considerable extent on the frequency. On short wavelengths a variation of as little as

a few hundred cycles may make an appreciable difference to fading. This is a serious difficulty with telephony reception, because it means that the modulation is continually being distorted, all the bass being missing at one instant and all the treble the next.

For Morse reception, however, this trouble does not arise, and fading may be countered to some extent by transmitting a modulated signal. Hence, when the carrier is fading the modulation may still be received reasonably well and vice versa. The method is not so effective as the spaced aerial arrangement just discussed, but it is, of course, much simpler.

## Ultra-Short Wave Aerials.

The types of receiving aerial systems employed for ultra-short waves are discussed in Chapter XII. They are generally similar to those employed for the transmission of these wavelengths, and constitute a special branch of the subject.

## Anti-interference Devices.

One of the principal troubles in modern reception, particularly in broadcast receiving, is that of local interference. Much modern electrical apparatus produces electrical disturbances. These take the form of very rapidly damped waves of various frequencies, and they have the same effect as a plain aerial, a form of transmitter which, as explained in Volume I, has been banned for many years owing to its very broad tuning and the interference which it causes.

The interference is transmitted in three ways—

(*a*) Direct radiation.
(*b*) Direct conduction.
(*c*) Mains radiation.

The first is self-explanatory. The interfering currents at the source generate rapidly damped waves which are radiated and picked up by the receiver. The range of such interference is limited to 10 or 20 yards, and this form of disturbance is most troublesome in industrial districts, though lifts and even household appliances such as refrigerators may be annoying.

Direct conduction is usually low-frequency in character and is only occasionally encountered in the case of sets operated from the electric mains. It usually arises from inadequate smoothing in the h.t. supply portion of the set or from bad layout or poor earthing. It is thus mainly a matter of design.

Ninety per cent of the interference is of class (*c*). The interference is conducted by the mains acting as a radio transmission line to the point where the receiver is located and is then re-radiated on to the receiving aerial. It may

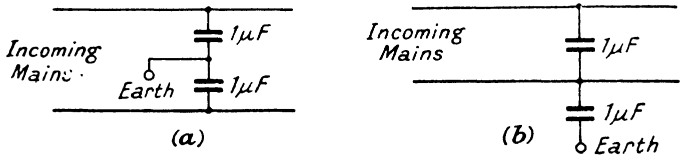



FIG. 83. SUPPRESSION CIRCUITS FOR FILTERING MAINS-BORNE INTERFERENCE

travel either symmetrically, going out on one main and returning on the other, or asymmetrically, going out on both mains together with the earth as a return. Both forms are encountered and the interference may be conducted in this way several miles from the source.

The remedy is to filter the mains at the point where they enter the building by connecting a condenser to earth from each line as shown in Fig. 83. The arrangement of Fig. 83(*a*) is used for d.c. mains, but on a.c. the centre point is live until it is earthed, since the condensers will pass a.c. through the body of the operative. The circuit of Fig. 83 (*b*) is thus recommended as the earth terminal is then always dead.

The most satisfactory remedy, of course, is to fit suitable suppressors at the actual source of the interference. The sparking at the brushes of a motor, for example, can often be prevented from causing interference by connecting two condensers across the brushes and earthing the centre point. The radio-frequency oscillations generated in the motor are then short-circuited to earth through the by-pass condensers and prevented from travelling along

the mains, which in the normal course of events act as aerials.

In other cases it is necessary to insert radio-frequency chokes in the leads as well as the bypass condensers, while of course all these methods only prevent the radiation of the disturbance from the mains and do not affect the direct radiation from the offending machine. The only remedy against this latter trouble is to enclose the whole machine in a shielding box of heavy gauge metal, and this is not always practicable.

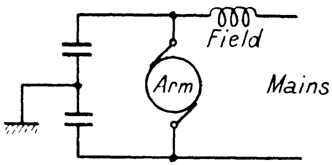



FIG. 84. SUPPRESSOR CIRCUIT FITTED TO MOTOR

The subject is one of some complexity and numerous special remedies are available for particular circuits. Electro-medical apparatus is very troublesome, for it actually uses radio-frequency currents and is thus a small transmitter of very bad noise. It can only be silenced by enclosing the equipment in an earthed screened room and filtering all outgoing mains and telephone wires inside the screen.

Trams and trolley buses are also troublesome, and it is becoming the practice to fit special h.f. chokes in the trolley arm leads and also in severe cases to connect condensers from the overhead lines to earth every hundred yards or so.

The reader is referred to a paper by Col. A. S. Angwin read before the Institution of Post Office Electrical Engineers, entitled "Interference with Wireless Reception Arising from the Operation of Electrical Plant," also to British Standard Specification No. 613, and to *Radio Interference Suppression* (Chapman & Hall) by the present author.

## Shielded Lead In.

Apart from the elimination of the disturbance at the source some measure of relief can be obtained by using a shielded down lead for the aerial. The aerial is provided with a short horizontal top and the down lead itself is

enclosed in a metal tube connected to earth. The fairly local direct radiation from interfering plant is thus prevented from affecting the greater part of the aerial, and, although there is some loss in the effective signal strength received from the required signal, the overall result is a marked improvement.

It is not usually practicable to employ an actual tube,

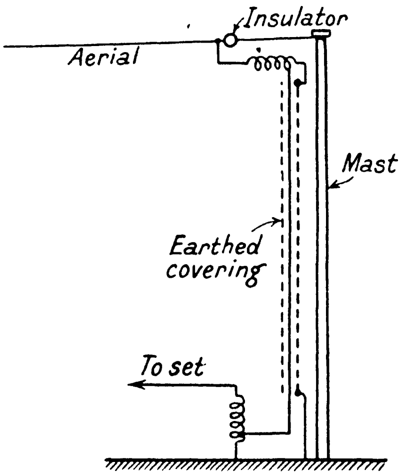



FIG. 85. SHIELDED LEAD-IN CABLE

so the lead is in the form of a cable having a fairly large outer diameter which is covered externally with a light metal braid. Even so, the capacity between this earthed outer covering and the down lead provides a short circuit for the signal-frequency currents, which are thereby prevented from reaching the receiver in their entirety.

To minimize this undesirable loss the lead-in is sometimes supplied with high-frequency transformers at both top and bottom of the aerial. The voltage picked up at the aerial head is stepped down through the first transformer, consisting usually of a simple tapped coil. The capacitance currents depend upon the voltage and are, therefore, much smaller than before. At the bottom of the aerial the voltage

is stepped up again through a similar transformer before being applied to the receiver. While there is some loss in this process it is usually less than that obtained by leakage if the transformers are not used, and it has in fact been found possible to employ cables of relatively small diameter for the lead-in, provided this impedance matching is used.

## EXAMPLES VIII

(1) Calculate the optimum aerial coupling for a transformer having a secondary inductance of 165 $\mu$H. and a resistance of 10 ohms for use with an aerial of 20 $\mu$H. and 200 $\mu\mu$F. at a frequency of 800 k/c.

What will be the step-up with this transformer at this frequency?

(2) Calculate the distance between the peaks of a band-pass tuner having $L_1 = L_2 = 150$ $\mu$H. and mutual inductance $= 3$ $\mu$H. at frequencies of 1 200, 900 and 600 k/c.

What would be the difference if the circuits were coupled through a common inductance of 3 $\mu$H.?

(3) In the circuit of (2) above replace the mutual inductance by a common capacitance as in Fig. 89 (*b*). What capacitance will be required to give the same peak separation at 900 k/c.?

What would be the peak separation at 1 200 and 600 k/c using this capacitance?

# CHAPTER IX

## COUPLED CIRCUITS

RADIO communication is very largely concerned with the transfer of energy from one circuit to another usually with one or both circuits tuned to resonance. It is, therefore, of importance to understand the mechanism of such energy transfer and to appreciate the reaction of one circuit on the other.

Reference has already been made in earlier chapters to this interaction and to the double-frequency effects known as double humping which can occur in certain circumstances. The complete theory of coupled circuits is apt to be involved and mathematical, but we shall consider here some of the simpler cases covering most of the circuits likely to be encountered in practice. The reader who wishes to investigate the subject more thoroughly is referred to *Electric Oscillations and Electric Waves* by Pierce or *High-frequency Alternating Currents* by McIlwain and Brainerd.

### Simple Inductive Coupling.

Let us consider first of all a simple mutually inductive arrangement as shown in Fig. 86. As explained in Volume I,

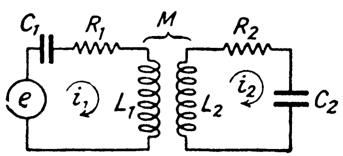



FIG. 86. SIMPLE MAGNETICALLY COUPLED TUNED CIRCUITS

Chapter VIII, the primary current induces a voltage in the secondary which produces a secondary current. This current in turn induces a voltage back into the primary in such a direction as to limit the primary current. Thus, the presence of the secondary causes an increase in the effective resistance of the primary as one might expect.

There is a similar effect on the reactance, but we must analyse the matter more closely to discover just how these effects operate.

From ordinary a.c. theory we can write

$$i_1(R_1 + jX_1) + jM\omega i_2 = e$$

$$i_2Z_2 + jM\omega i_1 = 0,$$ $Z_2$, being the impedance of the secondary, including the reactance of the transformer secondary.

Thus $i_2 = -jM\omega i_1/Z_2$.

Substituting this in the first expression we have

$$i_1(R_1 + jX_1) - j^2(M^2\omega^2/Z_2)i_1 = e$$

Therefore the effective primary impedance

$$Z_1' = e/i_1 = R_1 + jX_1 + M^2\omega^2/Z_2 \text{ since } j^2 = -1$$

But $\dfrac{M^2\omega^2}{Z_2} = \dfrac{M^2\omega^2}{R_2 + jX_2}$

Rationalizing this (see Appendix) we have

$$\frac{M^2\omega^2}{R_2 + jX_2} = \frac{M^2\omega^2(R_2 - jX_2)}{R_2^2 + X_2^2} = \frac{M^2\omega^2}{Z_2^2}(R_2 - jX_2)$$

Hence $Z_1' = R_1 + jX_1 + (M^2\omega^2/Z_2^2)(R_2 - jX_2)$

Separating resistive and reactive terms, we obtain

$$R_1' = R_1 + (M^2\omega^2/Z_2^2)R_2$$

$$X_1' = X_1 - (M^2\omega^2/Z_2^2)X_2.$$

From these basic expressions (which the reader will recognize as having appeared on several occasions in preceding chapters) we can calculate the primary current and estimate the behaviour of the whole network as far as concerns its impedance to the input voltage feeding the primary.

If we wish to calculate the secondary currents or voltages, we can then do so by calculating the voltage transferred according to the usual laws, *using for the primary current the modified value just determined*. Thus, in a simple inductively coupled circuit

$$e_2 = -jM\omega i_1 \quad \text{where } i_1 = e/(R_1' + jX_1')$$
$$i_2 = e_2/Z_2.$$

This follows directly from the fundamental expressions stipulated at the beginning of this section.

**Tuned Transformer.**

It will be seen that the primary resistance is always increased, but the reactance is decreased, unless $X_1$ and $X_2$ happen to be of opposite sign. Note also that since $X_2$ can be made zero by correct choice of capacitance or frequency, so can $X_1'$ by making $X_1 = (M^2\omega^2/Z_2^2)X_2$. The tuning point therefore will not occur with $X_2 = 0$, as it would do if the secondary were isolated, but a somewhat higher frequency which makes $X_2$ positive and equal to $X_1$. Apart from this, the circuit behaves in the same way as the secondary would have done by itself, giving a resonance curve of the usual form (except when the primary is tuned as we shall see shortly).

**Common Reactance and Capacitance Coupling.**

The mechanism is obviously similar if instead of a mutual inductance we have a common reactance of some sort. With a portion of the inductance common, as in Fig. 87($a$),

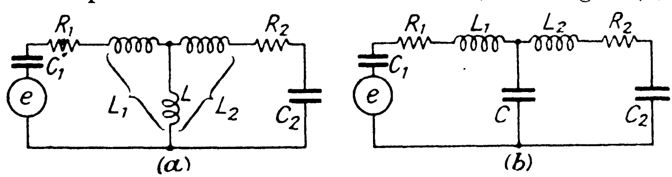



FIG. 87. ALTERNATIVE TYPES OF COUPLED CIRCUIT

for example, the expression $M\omega$ must be replaced by $L\omega$, $L$ being the common inductance. $L_1$ and $L_2$ will be the *total* primary and secondary inductances.

Alternatively we could use a capacitance, as in Fig. 87($b$), in which case the common reactance is $1/C\omega$ and the expressions become

$$R_1' = R_1 + R_2/C^2\omega^2Z_2^2$$
$$X_1' = X_1 - X_2/C^2\omega^2Z_2^2$$

## Tuned Primary.

An important practical case arises when the primary circuit is also tuned. The complete solution is complex but usually the two circuits are tuned to the same frequency which enables the theory to be simplified. Let us examine such a case and assume firstly that the coupling is small so that there is negligible interaction between the circuits.

As the frequency is varied through the resonant point the current in the primary increases to a maximum and falls off again according to the normal resonance curve. The voltage transferred to the secondary will rise and fall in the same manner and the secondary current will vary even more sharply, for the secondary reactance is varying at the same time.

Hence the secondary resonance curve is very sharp, being in fact the product of the normal curves for primary and secondary alone, but the actual peak value will only be relatively small.

We can increase the energy transfer by "tightening" the coupling, but the interaction between the circuits then becomes appreciable. The effective primary reactance $X_1' = X_1 - (M^2\omega^2/Z_2^2)\,X_2$, assuming mutual inductive coupling, and this varies with frequency in a peculiar manner just around the resonant point.

## Double Humping.

At resonance $X_1 = X_2 = 0$ so that $X_1'$ also $= 0$. The primary current, however, is not a maximum because the effective resistance has been increased by $(M^2\omega^2/Z_2^2)R_2 = M^2\omega^2/R_2$ since the secondary is tuned. Once the secondary goes off tune, however, $Z_2$ becomes appreciable and the additional load introduced by the secondary falls away rapidly, allowing the primary current to increase.

At the same time the effect of the secondary on the primary reactance is varying. If the frequency rises slightly, $X_1$ becomes a small positive value, *but so does* $X_2$, and since $X_1' = X_1 - (M^2\omega^2/Z_2^2)X_2$ the increase in $X_1'$ is checked. It is, in fact, possible for $X_1$ to equal $(M^2\omega^2/Z_2^2)X_2$ so that $X_1'$ again becomes zero and because of the reduction

of the effective primary resistance mentioned above, the current at this new resonant point may be greater than before. A similar effect occurs on the other side of the

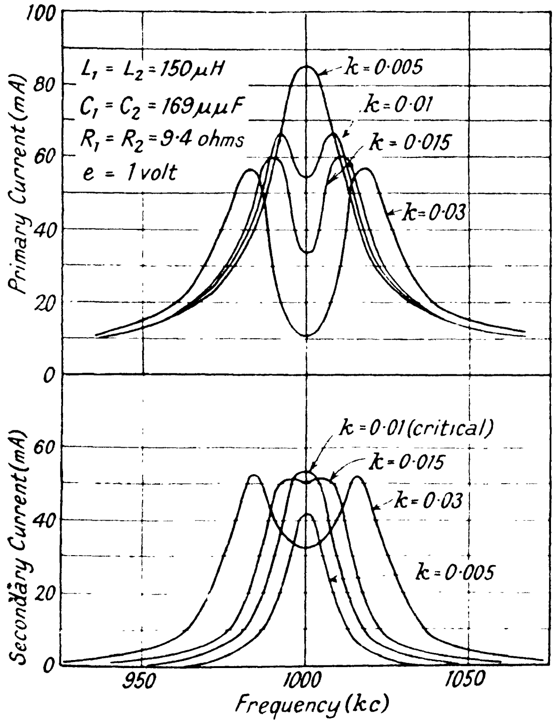



FIG. 88. PRIMARY AND SECONDARY RESONANCE CURVES FOR VARYING DEGREES OF COUPLING

true resonance point, so that the primary current is double-humped, showing two apparent tuning points as in Fig. 88.

**Secondary Current.**

The primary current $i_1 = e/Z_1'$

Hence the secondary current $i_2 = -jM\omega i_1/Z_2$

$$= -jM\omega e/Z_1'Z_2$$

If this expression is expanded by writing $Z_1'$ and $Z_2$ at length it becomes $-jM\omega e/[R_1R_2 - X_1X_2 + M^2\omega^2$

$$+ j(R_1X_2 + X_1R_2)]$$

At resonance $X_1 = X_2 = 0$ and the expression simplifies to

$$i_2 = -jM\omega e/(R_1R_2 + M^2\omega^2)$$

This is a maximum when the two terms in the denominator are equal so that $M^2\omega^2 = R_1R_2$, a condition which is termed *critical coupling* and has been mentioned in earlier chapters.

It can be shown that up to this critical coupling the secondary resonance curve only shows one tuning point. Beyond this value double humps appear in the secondary current. Moreover, the maximum value of the current, even at these peaks, is always less than the value with critical coupling, which represents the maximum possible. In practice the reduction is not serious at first and it may be desirable in particular circumstances to permit double humping to occur. The secondary peaks become pronounced at about twice critical coupling.

The maximum value of the secondary current is clearly $-jM\omega e/2R_1R_2$ (making $R_1R_2 = M^2\omega^2$) $= -je/2\sqrt{(R_1R_2)}$ (writing $\sqrt{(R_1R_2)}$ for $M\omega$). If $R_1 = R_2 = R$, this becomes, numerically, simply $e/2R$, i.e. one-half the current which would flow if the e.m.f. were applied to either circuit alone.

## Selectivity.

It has already been explained that the selectivity of the secondary with weak coupling is the product of that of the two circuits individually. As the coupling increases the resonance curve necessarily becomes broader. With critical coupling it is intermediate between this optimum value and that of the secondary alone and as the coupling becomes tighter it tends to a limit equal to that of the secondary alone while the top of the resonance curve shows marked double humps as already explained.

It is possible to determine the remote channel selectivity, i.e. the relative response several channels off tune by use of the coupled circuit laws, but the calculations become

complex and the calculation is of minor importance because in the design of a two-element or band-pass filter the main consideration is the shape of the resonance curve near the resonant point.

As already explained, the use of such circuits is adopted to provide a uniform response to a band of frequency immediately adjacent to the tuning point ($\pm$ 5 to 10 kcs.) so that the side bands of the transmission produced by the modulation shall be received at full strength and not attenuated as they would be by a normal resonant circuit. To calculate the shape of the top of the curves it is necessary to know the frequencies at which the peaks appear.

## Frequencies of Peaks.

The frequencies at which the circuits give maximum response are those at which the primary reactance is zero. In other words

$$X_1 - (M^2\omega^2/Z_2{}^2)X_2 = 0$$

It is found that this equation is not substantially affected by the secondary resistance. Hence we can make $R_2 = 0$, and if we then write $\omega = 1/\sqrt{(LC)}$ and $k$, the coupling factor, $= M/\sqrt{(L_1L_2)}$ we obtain the simple relation

$$f = f_o/\sqrt{(1 \pm k)}$$

where $f_o$ is the usual resonant frequency $= 1/2\pi\sqrt{(LC)}$ and $f$ is the frequency of the maximum response.

The shape of the curve depends both on the coupling factor and the $Q$ $(= L\omega/R)$ of the circuits. The better the circuit the more pronounced the humps and the object is to make the top of the curve as flat as possible. This is obviously equivalent to making the secondary current at the two tuning points equal to that at the true resonant frequency $f_o$ and by use of the equations already developed this can be shown to occur when $\sqrt{(Q_1Q_2)} = 1{\cdot}5/k$ where $Q_1 = L_1\omega/R$ and $Q_2 = L_2\omega/R_2$. The curve in Fig. 88 with $k = 0{\cdot}015$ shows the type of response obtained.

Under these conditions the voltage across the secondary is $e\sqrt{(Q_1Q_2)}/2$, i.e. one-half the voltage developed by an equivalent single circuit, which is a natural corollary from

the fact already mentioned that the secondary current with critical coupling is one-half the value with one circuit only.

The "pass band" of the filter is obviously the separation between the peaks which can be expressed as a fraction of the normal resonant frequency.

$$\frac{\text{Pass band}}{\text{Resonant frequency}} = \frac{\sqrt{(1+k)} - \sqrt{(1-k)}}{\sqrt{(1-k^2)}}$$

which, if $k$ is small, reduces simply to $k$.

For more detailed treatment the reader is referred to the articles in *Wireless Engineer* already quoted on page 145.

## Other Forms of Coupling.

The expressions so far developed apply to any form of circuit provided that the term $M\omega$ is replaced by an equivalent reactance. Fig. 89 shows four types of circuit in common use.

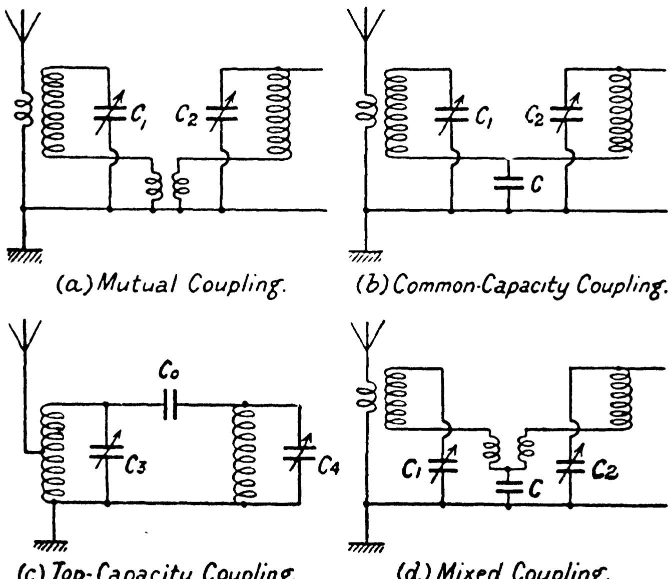



FIG. 89. TYPICAL BAND-PASS AERIAL COUPLINGS

Circuit (*a*) is a mutually coupled circuit of the type already discussed. The coils may be coupled in part as shown, or as a whole, as is usual in i.f. transformers, the coils being mounted in the same screening can be separated by a suitable distance to give the required coupling, usually critical coupling. Alternatively a common inductance may be used in which case $M\omega$ is replaced by $L\omega$. In all cases $L_1$ and $L_2$ must be the total inductances, including that of the coupling portions.

Circuit (*b*) is a common-capacitance coupled arrangement in which $M\omega$ is replaced by $1/C\omega$. An alternative form is shown in circuit (*c*). The coupling reactance here is not so easy to specify, but the circuit can be converted to the form of circuit (*a*) by using the transformations quoted in the Appendix. If this is done we arrive at the following—

$$C_1 = C'/C_4$$
$$C_2 = C'/C_3$$
$$C = C'/C_o$$

where $$C' = C_3C_4 + C_3C_o + C_oC_4.$$

## Free Oscillation.

Where the coupled circuits are not provided with a source of e.m.f., but are allowed to oscillate naturally, e.g. in a valve oscillator, similar effects occur with certain important reservations. If the coupling is weak, the energy in the primary is partially transferred to the secondary, the remainder being dissipated in primary losses.

As the coupling is increased the energy transfer increases to an optimum after which an unstable condition occurs in which the energy surges backwards and forwards from one circuit to the other as explained in Volume I, producing a complex double-frequency oscillation. The critical conditions to avoid this may best be arrived at by considering the frequencies at which the circuit will oscillate.

**The theory is simplified if we assume that both primary and secondary circuits are tuned to the same angular frequency $\omega$. Let us then assume that the two frequencies of the system are $\omega + \delta\omega$ and $\omega - \delta\omega$.**

At the upper frequency

$$X_1 = L_1(\omega + \delta\omega) - 1/C_1(\omega + \delta\omega)$$
$$= \frac{L_1C_1(\omega + \delta\omega)^2 - 1}{C_1(\omega + \delta\omega)}$$
$$= \frac{L_1C_1\omega^2 - 1 + 2L_1C_1\omega\delta\omega}{C_1(\omega + \delta\omega)},$$

neglecting the term involving $(\delta\omega)^2$, which is very small.

Now $L_1C_1\omega^2 = 1$, and since $\delta\omega$ is small $\omega + \delta\omega = \omega$ nearly. The expression then simplifies to $X_1 = 2L_1\delta\omega$.

Similarly, $X_2 = 2L_2\delta\omega$.

The primary reactance $X_1' = X_1 - (M^2\omega^2/Z_2^2)X_2$, and if the circuit is self-oscillating it will choose a frequency such that $X_1' = 0$.

$$\therefore \qquad 2L_1\delta\omega - \frac{M^2\omega^2 2L_2\delta\omega}{R_2^2 + 4L_2^2\delta\omega^2} = 0$$

i.e. $$8L_1L_2^2\delta\omega^3 + (2L_1R_2^2 - 2L_2M^2\omega^2)\delta\omega = 0$$

This has three roots, namely,

$$\delta\omega = 0 \text{ and } \delta\omega = \pm \tfrac{1}{2}\sqrt{\left[\frac{M^2\omega^2}{L_1L_2} - \frac{R_2^2}{L_2^2}\right]}$$

If the circuit is to be mono-oscillatory the second two roots must be imaginary, which means that the expression under the root sign must be negative. Hence

$$M^2\omega^2 < R_2^2L_1/L_2 \text{ or } M < (R_2/\omega L_2)\sqrt{(L_1L_2)}$$

which is the criterion quoted on page 8.

The volt-ampere criterion mentioned on page 25 follows from this and is sometimes a more convenient way of stating the conditions.

The existence of three possible frequencies may come as a surprise to the reader who is accustomed to consider a coupled circuit as having only two modes of oscillation. It should be noted, however, that if $M$ is greater than the critical value the frequency given by $\delta\omega = 0$ $(f = f_o)$ is unstable.

Suppose $f$ increases slightly. Then $X_2$ is no longer zero and $X_1'$ is reduced. The resonant frequency rises accordingly and the current rapidly slides into the upper stable frequency corresponding to the positive value of $\delta\omega$. Similarly if the frequency falls the circuit immediately takes up the lower value of oscillation, and stable conditions only obtain with one or other of these two modes.

# CHAPTER X

## FILTERS AND ATTENUATORS

It is often necessary in communication practice to cut off or attenuate certain frequencies while accepting others. We do this by passing the current through a network of inductances and capacitances arranged in suitable sequence. Such a system is known as a *filter*.

### Simple L-section.

The simplest type of filter is illustrated in Fig. 90. We are interested in the voltage on the output, which is

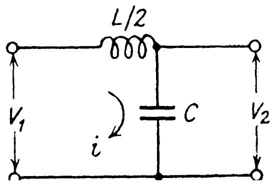



FIG. 90. SIMPLE L-SECTION FILTER

that developed across the condenser. Assuming we have no other load on the circuit the only current is that flowing through the inductance and capacitance in series. At low frequencies this will be small owing to the high reactance of the condenser. The reactance of the inductance will be small, and hence practically the full voltage will appear across the condenser.

As we raise the frequency the voltage drop on the inductance increases until we reach the point where the values of $L$ and $C$ resonate with the applied frequency. Here we have an equal distribution of the voltage, half being dropped on the inductance and the other half on the condenser, but due to the resonance the total reactance is reduced and the current increases accordingly. Hence, the voltage on the condenser can exceed the input voltage

just as in any ordinary resonant circuit, and we actually get a rise in the voltage at the end of the filter.

Beyond this point the voltage on the condenser begins to fall off rapidly and we obtain a cut-off, the actual characteristic being as shown in Fig. 91.

Now a simple L-section filter such as this is of limited use. For one thing, the cut-off is only gradual, and a

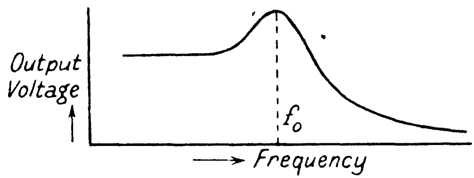



FIG. 91. TRANSMISSION CHARACTERISTICS OF FIG. 90 CIRCUIT

better filter can be obtained by connecting several sections in cascade. If this is done, we obtain a network rather like that of the transmission line discussed in Chapter III, and, for the reasons there explained, it is necessary to see that the various sections are suitably matched to one another and to the terminal load. Otherwise, reflection will occur, with consequent waste of energy, resulting in imperfect transmission.

## T- and $\pi$-sections.

Let us examine the impedance of a few typical filter sections. Fig. 92 shows a general case of a composite filter.

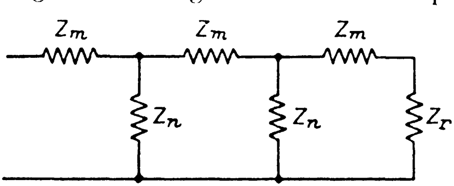



FIG. 92. GENERAL FILTER NETWORK

This can be split up into two types of section. One is the T-section as shown in Fig. 93($a$). Here the series impedance $Z_m$ is split into two, one half on each side. A number

of these sections placed in cascade will obviously be equivalent to the original network.

We can, if necessary, split the section into two half sections as at Fig. 93(*b*). To do this, the shunt impedance

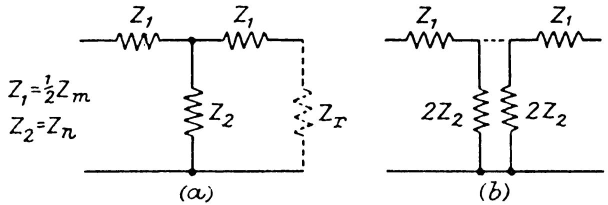



FIG. 93. T-SECTION FILTERS

is replaced by two impedances each twice as great, so that the two in parallel give the original value.

The second type is the $\pi$-section shown in Fig. 94. Here the series impedance is the full value, while the shunt impedances are each twice the original value. This network is, in fact, two half T-sections placed end to end. The

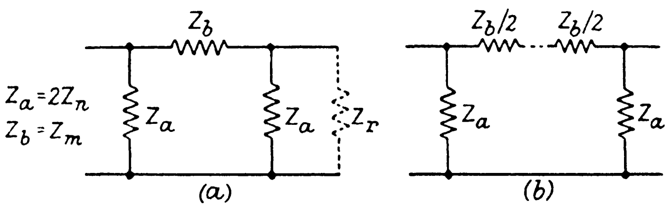



FIG. 94. $\pi$-SECTION FILTERS

$\pi$-section can also be split into two half sections which are identical with the T half-sections.

The T-section is also known as the *mid-series* termination and the $\pi$-section as the *mid-shunt* termination.

## General Filter Equations.

Let us now examine the impedance of these typical sections. For simplicity we will assume that the sections are symmetrical, as shown, but otherwise the impedances are unrestricted. Consider first the mid-series or T-section

(Fig. 93). We can write down the following expressions for the impedance of the filter viewed from the input end—

$$Z_o = \text{impedance with far end open } (Z_r = \infty)$$
$$= Z_1 + Z_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$
$$Z_s = \text{impedance with far end short-circuited } (Z_r = 0)$$
$$= Z_1 + \frac{Z_1 Z_2}{Z_1 + Z_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$
$$Z = \text{impedance with impedance } Z_r \text{ across output}$$
$$= Z_1 + \frac{Z_2(Z_1 + Z_r)}{Z_1 + Z_2 + Z_r} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

**Iterative Impedance.**

A case of particular importance arises if $Z = Z_r$ so that the input impedance is equal to the load or terminal impedance. Let $Z_k$ be the particular value of $Z_r$ which satisfies this condition. Then, from (3) above, we have

$$Z_k = Z_1 + \frac{Z_2(Z_1 + Z_k)}{Z_1 + Z_2 + Z_k}$$

$$\therefore Z_1Z_k + Z_2Z_k + Z_k^2 = Z_1^2 + Z_1Z_2 + Z_1Z_k + Z_1Z_2 + Z_2Z_k.$$

Cancelling out $Z_1Z_k$ and $Z_2Z_k$ on both sides we have

$$Z_k^2 = Z_1^2 + 2Z_1Z_2$$
$$\text{or} \quad Z_k = \sqrt{[Z_1(Z_1 + 2Z_2)]}.$$

This is known as the *iterative* or *surge impedance* of the filter and is important for the following reason: If the filter "looks like" $Z_r$ we can terminate it either with a load $Z_r$ or *by another filter section*, provided this is also terminated by the iterative impedance $Z_r$, and so on indefinitely. Thus, the condition for non-reflection in a multi-section filter is that each section shall be terminated with its correct iterative impedance.

It is worth noting, in passing, that $Z_k = \sqrt{Z_oZ_s}$, as the reader may verify for himself.

The iterative impedance may be written in another form which is sometimes useful. If $Z_m$ is the *total* series impedance ($= 2Z_1$) and $Z_n$ is the total shunt impedance ($= Z_2$), we can write

$$Z_k = \sqrt{[Z_mZ_n(1 + \rho)]} \quad \text{where } \rho = Z_m/4Z_n.$$

We shall see the advantage of this form in a moment.

## Mid-shunt Termination.

Let us now consider the mid-shunt or $\pi$-section. We treat this in the same way as before, and once again we can deduce the iterative impedance. This is found to be

$$Z_k' = \sqrt{\left[\frac{Z_a Z_b}{2 + Z_b/Z_a}\right]}$$

If again we write $Z_m$ for the total series impedance ($= Z_b$), and $Z_n$ for the *total* shunt impedance ($= Z_a/2$), we have

$$Z_k' = \sqrt{[Z_m Z_n/(1+\rho)]}.$$

Thus, the only difference between the mid-series and mid-shunt termination is in the position of the factor $(1+\rho)$, and these expressions for $Z_k$ in terms of $Z_m$ and $Z_n$ are easier to memorize accordingly.

## Cut-off Frequency.

The complete solution of a filter network involves mathematical treatment a little beyond the scope of this book, and a simplified treatment will be given. Consider a half-section mid-series terminated, which is equivalent to the simple L-section discussed at the beginning of the chapter (Fig. 90). The current through the section is

$$i = \frac{V_1}{j\omega L/2 + 1/j\omega C}.$$

The voltage across the output is the condenser voltage, which is

$$V_2 = \frac{i}{j\omega C} = \frac{V_1}{j\omega C(j\omega L/2 + 1/j\omega C)} = \frac{V_1}{1 - \omega^2 LC/2}.$$

Hence the ratio of $V_2$ to $V_1$ is

$$2/(2 - \omega^2 LC).$$

If $\omega^2 LC = 2$, this is infinity. This corresponds to the simple resonance previously mentioned, though in practice $V_2$ does not rise to infinity owing to the presence of resistance in the circuit. If $\omega^2 LC = 4$, the expression $= -1$, giving the same voltage at the output as on the input (but with a change of phase).

Beyond this the output voltage falls off continuously, and hence this point is called the *cut-off* point.

Now, we may rewrite this expression as follows—

$$\frac{-\omega^2 LC}{4} = -1 = \frac{j\omega L}{4/j\omega C} = \frac{Z_m}{4Z_n} = \rho.$$

Hence the cut-off point of a filter is obtained when $\rho = -1$.

This is not a rigid proof, but the fact is true and is in a convenient form. It applies, whatever the form of $Z_m$ and $Z_n$.

## Low-pass Filters.

We have seen that with this type of filter, voltage is transmitted with little attenuation up to the cut-off point, after which the output voltage falls off rapidly. Hence, this type of network is known as a *low-pass* filter. Let

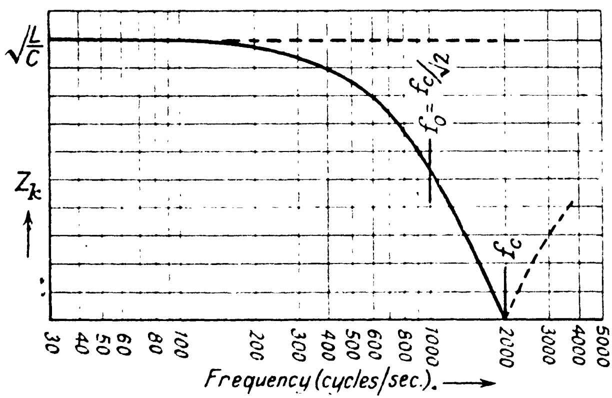



FIG. 95. VARIATION OF ITERATIVE IMPEDANCE WITH FREQUENCY (LOW-PASS FILTER)

us continue the analysis further. The *total* series impedance is $j\omega L$ and the shunt impedance $1/j\omega C$. The mid-series iterative impedance is thus

$$Z_k = \sqrt{\left[\frac{j\omega L}{j\omega C}\left(1 + \frac{j\omega L}{4/j\omega C}\right)\right]}$$ from the general expression given on page 168

$$= \sqrt{\left[\frac{L}{C}\left(1 - \frac{\omega^2 LC}{4}\right)\right]}.$$

Now the cut-off occurs when $\omega^2 LC = 4$, at which point $Z_k = 0$. When $\omega = 0$, $Z_k = \sqrt{(L/C)}$. Hence, over the "pass" range of the filter, the iterative impedance falls from $\sqrt{(L/C)}$ to 0, varying slowly at first and then rapidly as the cut-off frequency is approached; after which it becomes complex. We are, however, not interested in this region beyond the cut-off point.

Note that $Z_k$ has the dimensions of a resistance over the whole of the pass range, and the terminal load should, therefore, be non-reactive. The value of $Z_k'$, the mid-shunt iterative impedance, can be worked out by the reader. It will be found to vary between $\sqrt{(L/C)}$ and infinity over the pass range.

This variation of the iterative impedance is unavoidable with a simple filter, but over the majority of the pass range the value is not very different from $\sqrt{(L/C)}$, as shown in Fig. 95. Hence this value is usually used in filter calculations.

## Constant-K Filters.

If the product of the series and shunt impedances is constant at all frequencies so that $Z_m Z_n = K^2$, the filter is said to be a constant-$K$ type. The simple low-pass filter just described is of this type, and it should be noted that for this law to hold good the series and shunt impedances must be inverse, i.e. if one increases with frequency the other must decrease at the same rate.

With the simple low-pass filter, $Z_m = j\omega L$ and $Z_n = 1/j\omega C$, so that $Z_m Z_n = j\omega L \,.\, 1/j\omega C = L/C = R^2$, where $R$ is the iterative impedance. Hence, for a simple T-section filter terminated in a load $R$, we can define $L$ and $C$ exactly, as follows—

$$L = R/\pi f_c$$

and $C = 1/\pi f_c R$

where $f_c$ is the cut-off frequency.

## High-pass Filters.

If we wish to attenuate the low frequencies we use a *high-pass* filter as shown in Fig. 96. This works in just

the opposite manner from that just described. It will accept frequencies above a certain value, but will cut off below the critical frequency. Here, again, we can work out the action of the filter from simple resonant theory, or by making $\rho = -1$. Either method gives $4\omega^2LC = 1$, which corresponds to a frequency 0·7 times the resonant frequency $f_o$ (and not $1{\cdot}4f_o$ as before, which is obvious

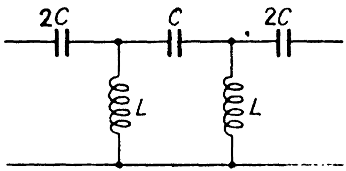



FIG. 96. HIGH-PASS FILTER

when one considers that we are running through resonance from a higher frequency this time).

The iterative impedance $Z_k$ is calculated as before and is found to be

$$Z_k = \sqrt{\left[\frac{L}{C}\left(1 - \frac{1}{4\omega^2LC}\right)\right]}.$$

At cut-off, when $4\omega^2LC = 1$, this is zero, while at frequencies *above* this (we are not interested in lower frequencies which are not in the pass range) the value rapidly rises and tends to a limiting value $\sqrt{(L/C)}$. This is similar to the low-pass case if the frequency is assumed to vary from infinity *down* to cut-off instead of from zero *up* to cut-off.

The filter is a constant-$K$ one, and if we put $Z_mZ_n = R^2$ we have

$$L = R/4\pi f_c$$

$$\text{and } C = 1/4\pi f_cR.$$

**Derived Filters.**

A series of sections of a simple filter network will provide a cut-off of increasing sharpness as the number of sections is increased. It is, however, possible to improve the cut-off with a smaller number of sections. For example, if

we include a tuned circuit in the shunt arm of the filter, such as is shown in Fig. 97(*a*), we shall obtain a point where the attenuation is complete, for when $L'$ and $C$ resonate, the voltage across the points $AB$ will be zero (assuming that the resistance in the network is negligible). On the other hand, beyond this point the voltage at the output will commence to rise and we shall obtain a characteristic somewhat similar to that shown in Fig. 97(*b*).

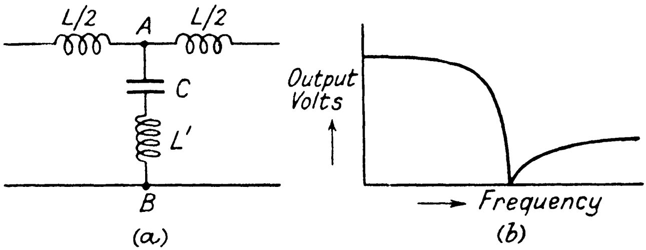



FIG. 97. DERIVED FILTER
By including an inductance in the shunt arm we can obtain complete cut-off

It is obviously desirable that this *derived* filter, as it is called, should have the same cut-off frequency and the same iterative impedance as the *prototype* from which it was obtained. Assuming a mid-series termination again, if we multiply $Z_1$ by a constant $m$, and divide $Z_2$ by the same constant, we have a filter having the same cut-off but different impedance. To comply with the latter condition, therefore, we make the additional impedance in the shunt arm equal to $Z_1(1 - m^2)/2m$, so that the total shunt impedance becomes

$$Z_2/m + Z_1(1 - m^2)/2m,$$ as shown in Fig. 98.

This, in conjunction with the modified series impedance $mZ_1$, will give the same value of $Z_k$ as for the simple filter, as the reader can verify for himself.

The additional impedance must, of course, be of the same type as $Z_1$ and the inverse of $Z_2$ (e.g. if $Z_2$ is a condenser the additional impedance is an inductance), and the frequency of maximum attenuation is given by the resonance condition between the two shunt impedances. This can

be shown to be $a$ times the cut-off frequency $f_c$, where $a = 1/\sqrt{(1 - m^2)}$ for a low-pass filter and $\sqrt{(1 - m^2)}$ for a high-pass filter.

Now $m$ may be any arbitrary number, but a case of special importance arises if $m = 0{\cdot}6$, for then the *mid-shunt*

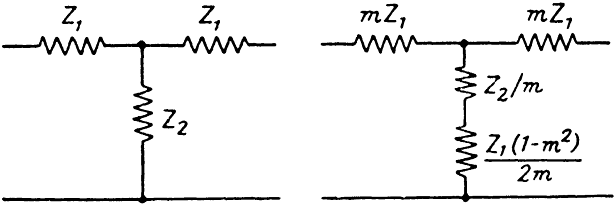



FIG. 98. MID-SERIES $m$-DERIVED FILTER SECTION

iterative impedance is constant and equal to $\sqrt{(L/C)}$ for the greater part of the pass band. When $m = 0{\cdot}6$, $a = 1{\cdot}25$ or $0{\cdot}8$ according to the type of filter, which results in a very sharp cut-off as shown in Fig. 101.

We can, if desired, obtain a derived filter from a mid-shunt or $\pi$-section as shown in Fig. 99. Here a similar

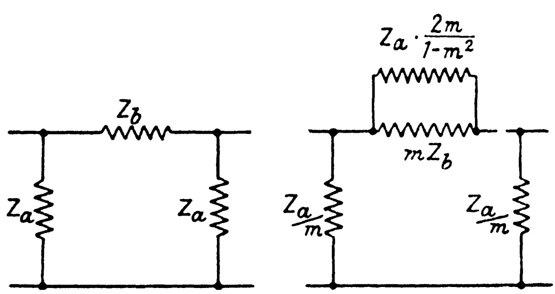



FIG. 99. MID-SHUNT $m$-DERIVED FILTER SECTION

procedure is adopted. We multiply the series impedance by $m$ and divide the shunt impedance by the same amount. Then we connect *across* the series impedance an inverse impedance $= [2m/(1 - m^2)]Z_a$. With this filter, if $m = 0{\cdot}6$, the *mid-series* iterative impedance is constant over the greater part of the pass band.

**Multi-section and Composite Filters.**

Filter sections of simple or derived types may be joined together provided they have corresponding terminations. Thus, a mid-series termination can be joined to another mid-series termination, but not to a mid-shunt termination.

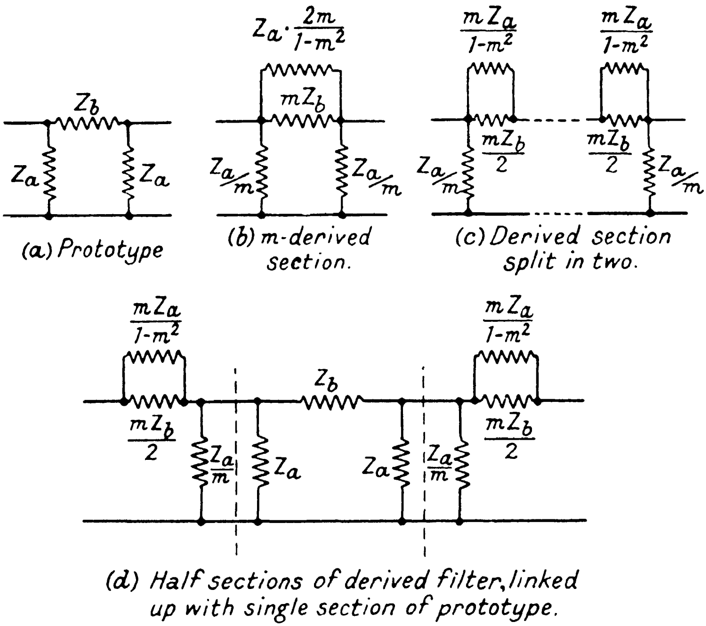



FIG. 100 ILLUSTRATING CONSTRUCTION OF COMPOSITE FILTER

Suppose now we wish to make up a filter consisting of one derived section and one simple section. Fig. 100 shows an example of how this may be done. First of all we derive a mid-shunt section. Then we split this into two half sections. Next, to the end of the derived section, we connect a simple mid-shunt or $\pi$-section, and to the end of this we connect the front half of the derived filter. Thus internally we have mid-shunt terminations adjacent

to one another, so avoiding reflection, while viewed as a whole the filter has mid-series termination, which as we have seen gives a practically constant iterative impedance.

Fig. 101 shows the response curve of a filter of this type. Curve $B$ is that of the derived filter giving a sharp cut-off but starting to pass frequencies above the maximum attenuation point. Curve $A$ shows the gradual cut-off of the

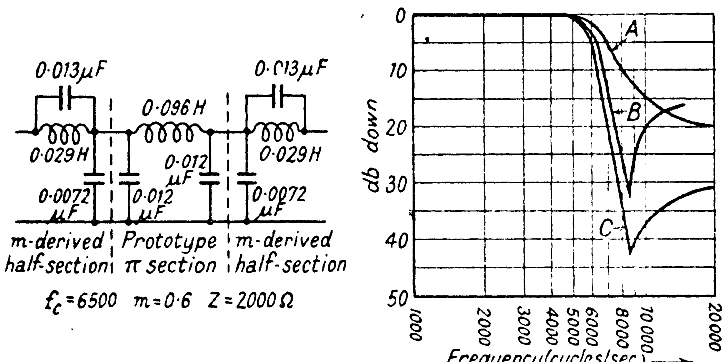



FIG. 101. TYPICAL COMPOSITE FILTER WITH TRANSMISSION CHARACTERISTICS

simple section with heavy attenuation later on. Curve $C$ is the combination of the two, giving a sharp cut-off and over 30 db. attenuation thereafter.

## Band-pass Filters.

Filters are sometimes required which will accept a band of frequency only. These are made up by using two filters in series, one a high-pass filter which is designed to cut off all frequencies below the lowest frequency in the band, and the second a low-pass filter which cuts off all frequencies above the highest desired frequency. Provided that the filters are suitably matched to one another so that their iterative impedances are equal, and equal to that of the input and the output circuit, frequencies within the band are transmitted practically unattenuated.

## Propagation Constant and Phase Angle.

The data given will enable practical filters to be designed, though the reasons for their behaviour have been somewhat briefly treated. It is difficult to discuss their behaviour more fully without recourse to mathematics. The theory is similar to that for a transmission line, as in Chapter III, except that $x$ is put equal to unity and the circuit values are those for the complete filter section.

Thus, for a filter correctly terminated with an impedance equal to the iterative impedance $v = V_0\varepsilon^{\gamma}$. The exponent $\gamma$ is the propagation constant and obviously $= \log_e v/V_0$. As we saw in Chapter III, it is not really constant since it is a complex quantity containing a resistive and reactive component, and hence varies with frequency.

If we write $\gamma = \alpha + j\beta$ it is found that over the pass range of the filter ($\rho = 0$ to $\rho = -1$) $\alpha$, the attenuation constant, is zero and $\beta$, the phase constant, is $2\sin^{-1}\sqrt{-\rho}$. Thus when $\rho = 0$ there is no phase shift while at the cut off frequency ($\rho = -1$) the phase shift is $2 \times \pi/2 = 180°$.

This is best understood by referring to the simple filter of Fig. 90. As we increase the frequency the output voltage gradually changes in phase, but remains of the same numerical value until the cut off frequency is reached when it has changed phase completely. Beyond this point the phase remains the same but the voltage begins to fall off.

Actually in Fig. 91 the voltage is shown rising to a peak just before the cut off, but this is because the filter is not correctly terminated. If a load resistance of the correct value is placed across the output, this resonant rise will be checked.

## Effect of Resistance.

In a practical filter, of course, the impedances are not pure reactances. The effect of resistance in the network is to cause some attenuation in the pass region, while the change of phase is delayed. At cut off it does not reach 180° and may never do so. Fig. 102 illustrates the effect for a single section in terms of $\phi$, the phase angle of the ratio $\rho = Z_m/4Z_n$.

With pure (inverse) reactances $\phi = 180°$, and under such ideal conditions the attenuation constant $\alpha$ is zero in the pass band (up to $\rho = -1$) and then rises sharply. Similarly the phase constant $\beta$ shows a gradual change in the phase of the output from zero to $\pi$ (complete reversal) after which

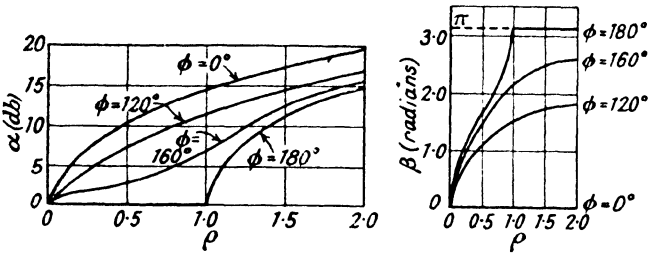



FIG. 102. THE EFFECT OF RESISTANCE ON THE ATTENUATION CONSTANT ($\alpha$) AND THE PHASE CONSTANT ($\beta$)
(The values of $\rho$ are all negative.)

there is no further phase shift but only attenuation as already mentioned.

The presence of even a small amount of resistance spoils this ideal action. The filter does not transmit signals without loss even in the pass band and as will be seen with $\phi = 160°$ the loss at the cutoff point is not zero but $7dB$—more than 2 : 1 down so that we shall only obtain half the correct output. Moreover, the cut off is no longer sharp but gradual, while as we introduce still more resistance conditions become steadily worse until when $\phi = 120$ there is no filter action but merely a progressive attenuation.

The phase shift is also reduced and continues to increase beyond the cut off point. When $\phi = 0$, corresponding to a simple resistance network, there is no phase shift at all and we merely obtain a fixed attenuation determined by the value of $\rho$.

It is clearly essential to reduce the resistances to a minimum. Even when $\phi$ is 160° we have largely destroyed the filter action. A phase angle of 160°, however, is quite poor. If we assume the resistance to be all in the inductance it is easy to show that $\phi = \pi + \tan^{-1} Q$, where $Q = L\omega/R$

On this basis $\phi = 160°$ is obtained with a $Q$ of 2·7 only. It is easy to obtain a $Q$ at least ten times as high, which makes $\phi = 178°$, under which conditions the performance will not be seriously different from the ideal. The figures in the pass band for values of $\phi$ near to 180° are shown in Fig. 103.

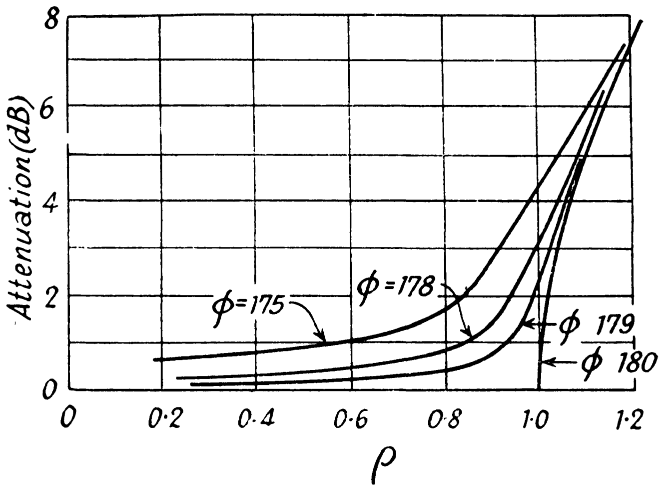



FIG. 103. ATTENUATION CURVES FOR FILTERS CONTAINING RESISTANCE

## Effect of Mismatching.

The performance both in and beyond the pass band is seriously affected if the matching is incorrect. A detailed discussion of this point is apt to become involved, though the calculations are simplified if we only consider symmetrical sections, as has been done here.

The curves of Figs. 102 and 103 are for a filter terminated in its correct (iterative) impedance, but we have seen (Fig 95) that this is not constant. $Z_k$ is actually zero at the cut off point so that no fixed value of resistance can give correct matching over the whole of the pass band.

Fig. 104 shows the effect of terminating the filter with various values of resistance. There are four curves as follows—

(1) Terminating resistance $P$ equal to the surge resistance $R\ (=\sqrt{(L/C)}\ )$. This gives negligible loss until $\rho = -0{\cdot}6$, after which a gradually increasing loss appears. There is no sharp transition at the cut off point $\rho = -1$, and the loss in the stop band is appreciably less than the ideal.

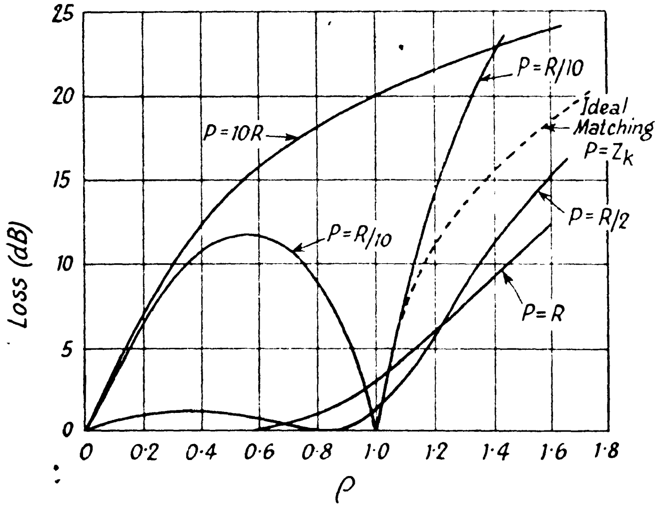



FIG. 104. EFFECT OF MATCHING ON PERFORMANCE

(2) $P = R/10$. Here there is perfect transmission at the cut off point and a very steep cut off thereafter, but there is a serious loss in the pass band.

(3) $P = R/2$. This is a compromise giving a small loss in the pass band, and a rather steeper cut off than with $P = R$.

(4) $P = 10R$. Here all trace of filter action has been lost.

The ideal matching curve is shown dotted for comparison. With $P = Z_k$ there is zero loss up to cut off and a rapidly rising loss thereafter.

All these curves apply to a mid series section, either low

or high pass, and show that the terminating impedance should not exceed $R$ and may often be made appreciably less with advantage. The effect of mismatching is reduced if several (correctly matched) sections are used in series because the intervening sections are then ideally terminated.

The tendency in practice is to design the filter for a cut off somewhat short of the theoretical point—e.g. if we require to accept frequencies up to 1 000 c/s we should design the filter with a cut off at about 1 300 c/s and use $P = R$ or possibly $R/2$.

For a mid-shunt section $Z_k$ varies between $\sqrt{(L/C)}$ and infinity over the pass band. Hence with such a filter inverse reasoning will apply so that we should tend to the use of higher loads than $R$. The $R/2$ curve will be obtained with $P = 2R$ (very nearly) and so on, so that with this type of section the filter action is destroyed if $R$ is too low.

With this simple treatment we must leave the subject. The reader desiring further information can refer to *High Frequency Alternating Currents*, by McIlwain and Brainerd, or to *Electric Wave Filters*, by F. Scowen.

## Equalizers.

An equalizer is a network intended to counteract irregularities in supply or arising from an uneven amplification in the circuit. A gramophone pick-up, for example, gives better response at some frequencies than others, but this could be corrected by connecting an equalizer across the circuit which accentuated the frequencies in which the pick-up was deficient and vice versa. Similarly, an amplifier or transmission line may have uneven characteristics, and these may be compensated by the use of suitable equalizers.

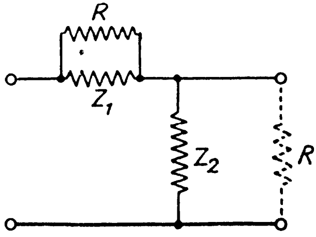



FIG. 105. EQUALIZER NETWORK

A very simple form of equalizer is obtained by connecting a resonant circuit across the supply. This would absorb frequencies at and around the resonant value. The extent

of the absorption could be controlled by including a series resistance at the circuit.

The design of equalizers is simplified if the impedance of the network can be made purely resistive at all frequencies, and such an arrangement is termed a *constant-resistance* equalizer. Fig. 105 shows a network feeding into a resistance $R$ and having an equal resistance $R$ across $Z_1$. If $Z_1Z_2$ equals $R^2$, the whole device has a resistive impedance equal to $R$.

This is easily proved as follows—

$$\text{Total impedance} = \frac{Z_1R}{Z_1 + R} + \frac{Z_2R}{Z_2 + R}$$

Writing $Z_2 = R^2/Z_1$ we have

$$Z = \frac{Z_1R}{Z_1 + R} + \frac{R^3}{Z_1(R^2/Z_1 + R)}$$

$$= \frac{Z_1R}{Z_1 + R} + \frac{R^2}{Z_1 + R} = \frac{R(Z_1 + R)}{Z_1 + R} = R$$

It follows from this condition $Z_1Z_2 = R^2$ that $Z_1$ and $Z_2$ must be inverse, i.e. if one is inductive the other must be capacitative. If they are alike their product is still reactive. Also a series resistance in one impedance must be represented by a parallel resistance in the other.

Let $Z_1$ be an inductance with a resistance in series so that $Z_1 = r_1 + j\omega L$. $Z_2$ will then be a condenser with a resistance in parallel, so that $Z_2 = 1/(1/r_2 + j\omega C)$.

$$\text{Then}\quad Z_1Z_2 = \frac{(r_1 + j\omega L)}{1/r_2 + j\omega C} = \frac{r_1r_2(1 + j\omega L/r_1)}{1 + j\omega Cr_2}$$

If we make $L/r_1 = Cr_2$ (i.e. $r_1r_2 = L/C$), this reduces simply to $r_1r_2$. Hence

$$Z_1Z_2 = r_1r_2 = L/C = R^2$$

A similar calculation for an inductance with a parallel resistance and a condenser with a series resistance produces the same result, so that the expression is general.

**Attenuation.**

Which form we use depends upon the conditions to be fulfilled. The attenuation of the network is given by the

ratio of the total impedance (which is $R$) to that of $Z_2$ and $R$ in parallel.

$$\therefore \ \alpha = R/[Z_2R/(Z_2 + R)] = (Z_2 + R)/Z_2$$

Alternatively since $Z_1Z_2 = R^2$ we can write

$$\alpha = (R + Z_1)/R$$

which may be a more convenient form.

Let us take as an example a network to give a uniform characteristic above 500 c/s, but a rising characteristic below this frequency, such that the output at 50 c/s is four times that at 500 c/s. This can be done by making $Z_2$ a condenser with a resistance in series. At low frequencies the condenser will have a high reactance, so that $Z_2$ is large and little attenuation occurs.

As the frequency rises the condenser reactance falls so that increasing attenuation is produced until the condenser reactance becomes small compared with the series resistance $r_2$. From this point on the attenuation is constant and determined solely by $r_2$.

A certain amount of trial and error is now required for we must assume some initial attenuation at 500 c/s. Otherwise $Z_2$ at this frequency would require to be infinite. Let us assume a 25 per cent loss so that $\alpha_1 = 1{\cdot}25$.

Then $\alpha_2 = 4 \times 1{\cdot}25 = 5$. But we have said that at 500 c/s the attenuation is to be determined solely by $r_2$, so that

$$\alpha_2 = 5 = (50\,000 + r_2)/r_2 \text{ whence } r_2 = 12\,500$$

At 50 c/s on the other hand, $Z_2$ is reactive, and we must use the full expression for $\alpha$.

$$\begin{aligned}\alpha_1 = 1{\cdot}25 &= (Z_2 + R)/Z_2 \\ &= [(R + r_2) + 1/j\omega C]/(r_2 + 1/j\omega C)\end{aligned}$$

We shall not be seriously in error if we neglect $r_2$ in comparison with $1/j\omega C$ in the denominator, and the expression then simplifies to

$$\begin{aligned}\alpha_1 = 1{\cdot}25 &= 1 + (R + r_2)j\omega C \\ \therefore \ \alpha_1^2 = 1{\cdot}625 &= 1 + (R + r_2)^2\omega^2C^2\end{aligned}$$

Whence, with $\omega = 314$, $C = 0{\cdot}04\mu F$.

For the series arm we must use an inductance with a resistance $r_1$ in parallel both determined from the relation

$r_1 r_2 = L/C = R_2$, whence $L = 99H$ and $r_1 = 200\,000$ ohms, the complete network being as shown in Fig. 98.

It should be checked that the reactance of $C$ is actually small compared with $r_2$ at 500 c/s. It will, in fact, be 8 000 ohms which is not really negligible, and the attenuation at 500 c/s would be 4·35 instead of 5. In other words, the rising characteristic would start somewhat above 500 c/s, and the output voltage would be about 18 per cent up at 500 c/s.

This is actually quite a good compromise. We could obtain a better ratio of max. to min. attenuation, but only

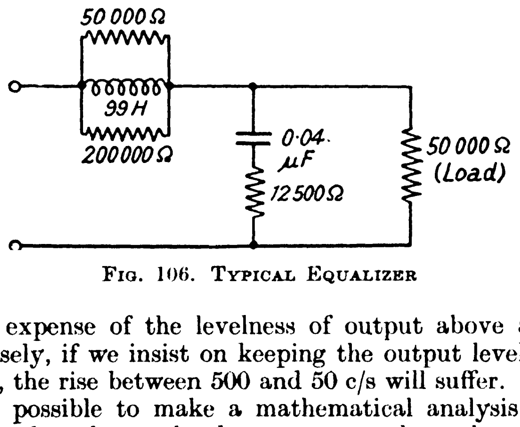



FIG. 106. TYPICAL EQUALIZER

at the expense of the levelness of output above 500 c/s. Conversely, if we insist on keeping the output level above 500 c/s, the rise between 500 and 50 c/s will suffer.

It is possible to make a mathematical analysis of the problem, but the results do not present themselves in convenient form, and intelligent trial and error applied to the principles already given will provide satisfactory practical results.

It is, perhaps, desirable to make sure that the problem is practicable before starting the calculations. An approximate check may be made as follows—

In whichever impedance has its elements in series ($Z_2$ in the case above), let $p_1$ and $p_2$ be the ratios of reactance to resistance at the frequencies in question. The maximum ratio of the two attenuations which can be expected is then $p_2/\sqrt{(1 + p_1^2)}$, and in general it will be somewhat less than this.

It should be noted, in conclusion, that $r_L$, the resistance of the inductance in $Z_1$, should be kept reasonably small. It should strictly be compensated by a resistance $R^2/r_L$ across the whole of $Z_2$, but a trial calculation shows that this is negligibly high provided $r_L$ is only 1 000 ohms or so.

**Phase Shift.**

Although the equalizer behaves as a resistance $= R$ the phase of the output voltage is not the same as the input voltage. The phase shift produced, however, is constant

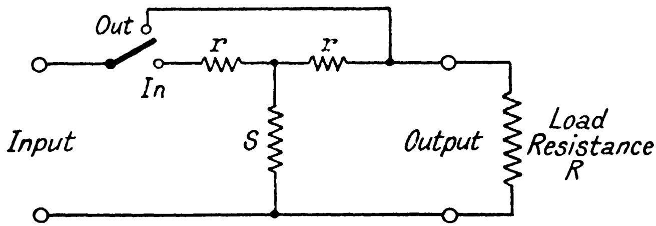



FIG. 107A. SIMPLE T-SECTION ATTENUATOR

at all frequencies if the network has been correctly designed which is an important feature not possessed by simpler networks.

This may be proved by developing the expressions for the series and shunt impedances and comparing the reactive terms which will be found equal at all frequencies provided that $L/C = R^2$.

**Attenuators.**

Finally, we come to the consideration of attenuator networks. These are arrangements of resistances designed to cut down the voltage by a known and constant amount irrespective of frequency. This is usually done by a T-section network as shown in Fig. 107A. The requirement is that the impedance of the attenuator shall always be the same and equal to load resistance.

The attenuator is thus a special case of a T-section filter terminated in the iterative impedance. The load resistance $R$ must be $\sqrt{[r(r + 2s)]}$.

The attenuation of a given section is easily deduced. Let $P$ be the resistance of $s$ in parallel with $R + r$. We have across the input a potentiometer made up of $r$ and $P$ in series. Across $P$ is a further potentiometer consisting of $r$ and $R$ in series. Hence, the output voltage

$$V_2 = \frac{P}{P + r} \cdot \frac{R}{R + r} V_1$$

Hence $\dfrac{V_2}{V_1} = k = \dfrac{PR}{(P + r)(R + r)}$.

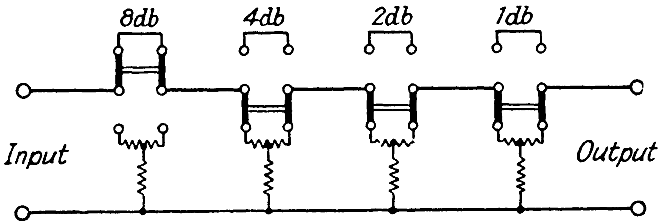



FIG. 107B. T-PAD ATTENUATOR COVERING A VARIATION OF 1 to 15 DB SET TO GIVE 7 DB ATTENUATION

If the network is terminated in its iterative impedance, the input resistance $= P + r = R$, so that $P = R - r$

Hence $k = (R - r)/(R + r)$.

This expression and that previously given for the iterative impedance completely determine $r$ and $s$. By rewriting the equations we have

$$r = R\left(\frac{1 - k}{1 + k}\right)$$

$$\text{and } s = 2R\left(\frac{k}{1 - k^2}\right)$$

Provided the resistances obey the laws given above, the effective resistance is always the same, and several sections, each having different attenuation factors, may be connected in series. Fig. 107B shows a network capable of giving 15 db attenuation in steps of 1 db.

## Variable Attenuators.

Where correct matching is required an attenuator of the form of Fig. 107 is essential. Fine adjustment is obtained, if required, by adding further sections—e.g. 0·4, 0·4, 0·2, and 0·1 db to provide fractions of 1 db.

Occasions arise, however, when approximate matching is sufficient, as with an oscillator or amplifier which is designed to feed into a specified load. Here the performance will not suffer seriously if the impedance varies by 10 or even 20 per cent. In such cases it is possible to make up attenuators having a continuously variable section as shown in Fig. 108. This is in the form of a tapped variable resistance, the taps being connected to earth through shunt arms so that the device is equivalent to a series of T-pads with the input taken to a variable point along the network.

It is impossible, with such a network, to fulfil the condition that the attenuator shall look like $R$ when terminated with $R$ and some compromise is therefore necessary. A commercial form of attenuator of this type is designed to look like 450 ohms within 10 per cent when terminated in 600 ohms, and a fixed resistance of 150 ohms is introduced to bring the input impedance up to 600 ohms. This introduces a permanent loss of 2·5 db, but beyond this point a continuous variation of output is obtained with a 20 db variable section and a number of fixed 20 db pads totalling 100 db in all.

## Ladder Attenuators.

In r.f. practice use is made of attenuators which are required to work into a high impedance. Such a case occurs in the output of a signal generator which works into a low resistance of 10 or 20 ohms, but which feeds the tuned aerial circuit of a receiver or the grid circuit of a valve, either of which, when tuned, has an impedance of some thousands of ohms which is negligibly high compared with the attenuator impedance.

For such work a ladder attenuator of the form of Fig. 109 is used. The input, fed in at the point A, is progressively attenuated in fixed steps, the output being taken from the points $A$, $B$, $C$, or $D$. If the impedance of the output is

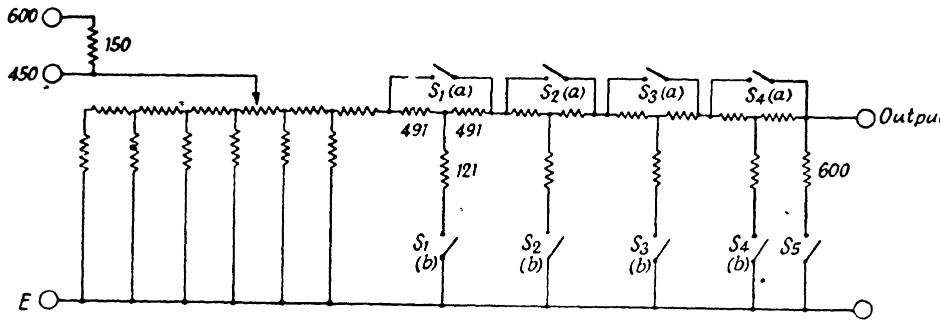



FIG. 108. EXAMPLE OF ATTENUATOR HAVING A CONTINUOUSLY VARIABLE SECTION ASSOCIATED WITH A SERIES OF FIXED SECTIONS

$S_1$ to $S_4$ are two-contact switches such that when (*a*) is open (*b*) is closed, and vice versa

high the input resistance of the network is clearly not affected by the position of output tap. Moreover, it is simple to arrange that the resistance of the network viewed from the output end is always the same (and equal to the input resistance) at any of the tapping points, so that the device behaves like a generator of constant source impedance.

Consider first the end section alone. The resistance to

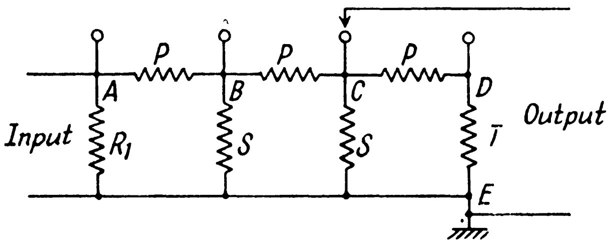



FIG. 109. LADDER ATTENUATOR NETWORK

the right of $C$ is $P + T$. If this step is to attenuate $n$ times $(P + T) = nT$.

To the right of $B$ the resistance is $P + Z$ where $Z$ is the resistance of $S$ in parallel with $P + T$. But if we are to maintain the same attenuation per stage $Z$ must be equal to $T$.

Hence $Z = T = nTS/(nT + S)$

whence $$S = nT/(n - 1).$$

Similar reasoning applies as we move progressively to the left, for the network to the right of $B$ can be replaced by a simple end section $P + T$ which gives the same value for the shunt resistance at $B$, and again to the right of $A$ we can replace the whole network by $P + T$.

The input resistance of the whole network is thus $R_1$ in parallel with $P + T$. If we make $R_1 = T$ the whole network becomes symmetrical. The resistance at any of the points $A$, $B$, $C$, or $D$ then becomes $(P + T)T/(P + T + T)$ $= nT^2/(nT + T)$

$$= nT/(n + 1) = R$$

Hence $$T = (n + 1)R/n$$

If $n = 10$ and $R = 10$, $T = 11$ ohms, $S = 12{\cdot}2$, and $P = 99$ ohms, which are values commonly used in r.f. attenuator networks.

If $R_1$ is not made equal to $T$ the input resistance will be $R_1(P + T)/(R_1 + P + T)$. This will still remain constant whatever the position of the output tap, but the resistance at the points $B$, $C$ and $D$ will not be the same. If this point

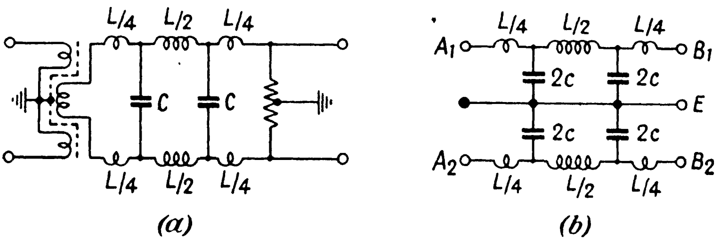



FIG. 110. EXAMPLES OF BALANCED NETWORKS

(a) Low-pass filter with screened and balanced input transformer and balanced transformer load.
(b) Symmetrical low-pass filter without terminal impedances.

is not of importance it may be found convenient to adopt a value of $R_1$ different from $T$.

## Balanced Networks.

It has already been explained in Chapter III that the use of an unbalanced line (i.e. with one side earthy) may give rise to cross talk, so that symmetrical arrangements are often employed. An attenuator or filter for use with such a system must be similarly "balanced."

This is achieved by splitting the series impedances in two and inserting one half in each line as shown in Fig. 110 (a). The calculations are as before, the two separate elements being considered as one. If the network includes a terminating impedance or transformer this will be centre-tapped to provide the earth point. Such terminating devices are themselves symmetrically constructed to equalize the stray capacitances. An earth screen is often provided in addition, the device being known as a screened and balanced transformer.

If no such termination is provided the shunt impedances must be similarly split into two elements in series, as shown in Fig. 110 (*b*).

## EXAMPLES X

(1) Calculate the resistances of a T-section attenuator to give a drop of 20 db. ($V_2/V_1 = 0{\cdot}1$) working into a 600-ohm load.

(2) Design an equalizer working into a load of 20 000 ohms to pass 10 per cent of the voltage up to 250 cycles and 90 per cent at 5 000 cycles.

(3) Design a composite filter comprising one simple section and one derived section to comply with the following conditions—

Cut off frequency = 6 500 cycles per sec.
Load = 2 000 ohms.

Make $m = 0{\cdot}6$ and use a mid-shunt derived section.

(4) Taking a general network of the type shown in Fig. 105 prove that—

(*a*) The input resistance = $R$;

(*b*) The attenuation is identical with that produced by the series arm only acting in series with $R$ provided that $Z_1Z_2 = R^2$.

# CHAPTER XI

## COMMERCIAL RADIO TELEPHONY

POINT to point radio telephone work has certain features of interest over and above the normal technique of radio telephony. A consideration which weighed considerably in the earlier days, but is not so important with modern short-wave transmission, is that of the power radiated. In an ordinary modulated wave 70 per cent of the power is in the carrier, which is, of course, useless as regards the

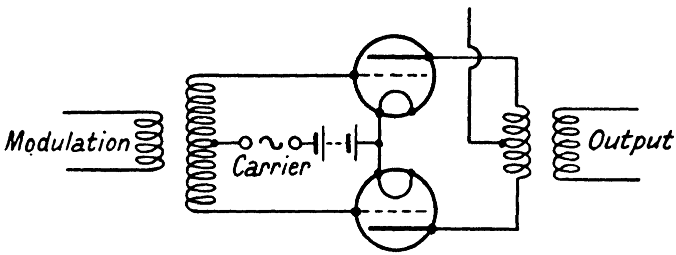



FIG. 111. CARSON BALANCED MODULATOR CIRCUIT

actual communication of intelligence, and the remaining 30 per cent is divided equally between the two side-bands. Attempts were made quite early, therefore, to eliminate the carrier wave.

Fig. 111 shows a circuit for doing this. A push-pull arrangement is employed and the carrier and modulating voltage are introduced separately. The former is introduced so as to affect the grids of the two valves in the same phase, whereas the modulation is introduced in the normal push-pull manner, so that the grids of the valves are in opposite phase. The usual centre-tapped output transformer is used, with the result that the carrier voltage does not appear in the secondary, whereas the modulation combines with the carrier to give side-bands in the ordinary way.

At the receiving end it is necessary to reintroduce the carrier wave, which is done by a local oscillator of the correct frequency, amplitude, *and phase*. This last requirement is most important for the correct modulation depth is only preserved if the relative phases of the carrier and side bands are maintained.

In practice this requirement demands such a highly elaborate technique that carrier suppression is rarely used.

## Single Side-band Working.

It is possible, however, to obtain similar and rather better results by filtering out one of the side bands after the carrier has been suppressed. This is done by a system of high- and low-pass filters of the type described in the preceding chapter, and results in the suppression of everything but the one set of side frequencies. This contains all the information necessary to communicate the intelligence, and if this is recombined with a carrier at the receiving end speech is obtained once again. The frequency of the re-introduced carrier must be as accurate as possible and one method adopted is to transmit with the side bands a fraction of the original carrier which acts as a "pilot carrier' and is used at the receiving end to synchronize a local oscillator.

Eckersley has suggested that a system using the carrier, one complete side band, and about 250 c/s of the other could be used without modification to the receiver and would result in less interference between stations, but the suggestion was never adopted.

The reader who is interested should refer to a paper by P. P. Eckersley entitled "Asymmetric Side-band Broadcast Transmissions," *Journal I.E.E.*, Vol. 77, p. 517, and also to *Short Wave Radio* by the present author, where the subject is discussed in detail.

Single side-band working is used considerably on carrier-current telephony. This is a radio-frequency transmission superposed on ordinary land lines. One line may be made to carry several transmissions of different frequencies in addition to low-frequency telephone or telegraph signals.

## Power in Modulated Wave.

A modulated wave can be represented by

$$\begin{aligned} v &= (A + B \sin pt) \sin \omega t \\ &= A \sin \omega t + B \sin pt \sin \omega t \\ &= A \sin \omega t + \frac{B}{2} \sin(\omega - p)t - \frac{B}{2} \sin(\omega + p)t \end{aligned}$$

which discloses the existence of the carrier and the two side bands.

The power radiated is proportional to

$$A^2 + 2\left(\frac{B}{2}\right)^2 = A^2 + \frac{B^2}{2}$$

Thus if $B = A$ (100 per cent modulation), the power is increased 50 per cent but only one-third of the power is in the side bands, two-thirds being wasted in the carrier.

If we suppress the carrier, therefore, there is an immediate saving of two-thirds of the power, while conversely with the same power we have a gain of 3 : 1 or 4 db.

The same applies to single side-band working, for although we save half the side-band power by suppressing one of the side bands we find that the received field strength is halved and that to obtain equivalent strength the amplitude of the transmitted side band has to be doubled.

Still more striking power economy results if the modulation is less than 100 per cent, as is usually the case in practice. The ratio of power expended without and with carrier is $B^2/(A^2 + B^2/2) = \frac{2m^2}{2 + m^2}$ where $m$ is the depth of modulation $= B/A$. If $m = 0.1$ only, the power ratio $= 0{\cdot}01$ nearly, giving a 100 : 1 improvement.

At the receiver the received field strength increases as the square root of the power ratio giving a 2 db. improvement for 100 per cent modulation.

## Ring Modulator.

An alternative form of modulator which suppresses both the carrier and the modulation frequency, leaving only the

side bands, is the ring modulator shown in Fig. 112. The carrier is impressed across the points $PQ$ and flows through the two halves of the transformer windings in opposition. The positive half cycles flow through the rectifiers $A$, $B$ and the negative half cycles through $C$, $D$.

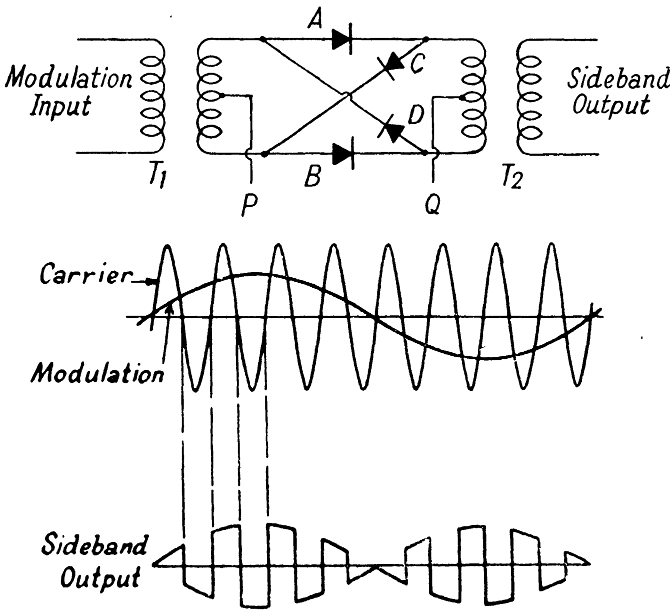



FIG. 112. RING MODULATOR CIRCUIT

While the rectifiers $A$, $B$ are working, $C$ and $D$ will be subject to a negative voltage and therefore will remain non-conducting even if further voltages (e.g. modulation voltages) are impressed across them, provided these additional voltages do not exceed the carrier voltage. Similarly when the rectifiers $C$, $D$ are operative $A$ and $B$ are cut off. No carrier voltage appears in the output due to the symmetry of the arrangement.

The modulation is impressed across $T_1$ and current flows

alternately through $A$, $B$ and $C$, $D$ producing side frequencies in $T_2$ by combination with the carrier. These side tones do not cancel out so that the desired side bands appear in the output while the carrier is suppressed.

The arrangement, in fact, is a switch of which the connexions are changed over by the carrier, provided that the carrier voltage exceeds the modulation voltage (as it always does in practice). Fig. 112 shows this operation physically, indicating the form of the side band output obtained in the secondary of $T_2$.

Centre-tapped resistances may be used in place of $T_1$ and $T_2$.

## Inversion and Scrambling.

Privacy is one of the principal requirements of a commercial telephone service, and the fact that ordinary broadcast telephone signals could be easily picked up militated at first against the use of radio telephony. The difficulty has largely been overcome with modern short-wave circuits which employ beam transmission and reception, while to make doubly sure the speech is often inverted.

For this purpose the speech frequencies before modulation are combined with a fairly high frequency producing side bands in the ordinary course of events. The lower side band only is chosen, so that the lowest speech frequency is now the highest modulation frequency. These frequencies are now heterodyned with another frequency lower than the lowest frequency in the band. The difference tone is used, all other frequencies being filtered out, and we are left with a modulation much the same as the original as regards range, but completely inverted.

Such a modulation is quite unintelligible to the ordinary listener, and privacy is thus assured. It is translated at the receiving end by going through the reverse processes, in the requisite order.

As a typical example we may take a frequency range of 100–3 000 c/s. This would be arranged to modulate an oscillator of, say, 50 kc/s. The lower side band, ranging from 47 000 to 49 900 c/s, would be selected, and subsequently made to heterodyne another frequency to reduce

the actual order of frequency. This is often done in two stages. For example, the first stage would use an oscillator of 42 000 c/s giving difference frequencies of 5 000 to 7 900 c/s, the other frequencies being suppressed, and this if necessary could again be heterodyned with an oscillator of 4 900 c/s which would give the original scale of frequencies ranging from 100 to 3 000 c/s, but completely inverted.

Another method is to transpose certain bands of frequency, which is done by selecting the required bands with filters and heterodyning them to different degrees so that when recombined they occur in a different order.

In practice both methods are employed either separately or together and speech broken up in this general manner is said to be *scrambled*.

## Land Line Connexion.

A special circuit is necessary to link a radio channel to a land line. The radio channel is a four-wire system having a pair of leads for the transmitted speech and another pair for the received signals. These are combined in a special transformer known as a *hybrid coil*, which enables both signals to be carried over the same pair of wires as is done with ordinary telephony.

Fig. 113 illustrates the system. Speech from the subscriber induces voltages in windings $A$ and $C$. The former passes to the transmitting repeater (which is merely an amplifier) and this in turn modulates the transmitter. The voltage in winding $C$ can produce no effect, because it is in the output of the receiving repeater and cannot force its way back owing to the undirectional characteristic of any valve amplifier.

Speech picked up on the receiver is amplified by the receiving repeater and appears in windings $C$ and $D$. Currents are induced in the land line, where they travel on to the subscriber, and also in the balancing network, which is an artificial line having the same characteristics as the actual line. These currents also induce voltages in $A$ and $B$, but as one of the transformers is reversed the

voltages are in opposition and cancel out. Thus, the received speech is prevented from operating the transmitter.

**Singing.**

Any interaction between the transmitter and receiver will give rise to a continuous oscillation or *sing*. There are two forms encountered in practice, one being a half sing when energy is fed back from the local transmitter

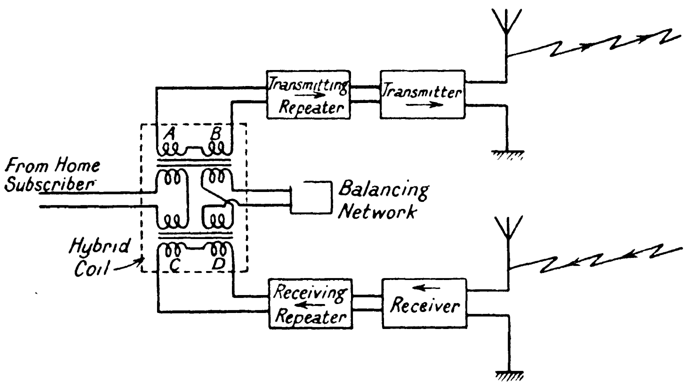



FIG. 113. ILLUSTRATING CONNEXION OF RADIO TO LAND-LINE

to the local receiver, while the other is an all-round sing from the local transmitter to the distant receiver and back through the distant transmitter to the local receiver.

The first difficulty can be overcome by adequate design of the local receiver. Given sufficient wavelength separation and a selective circuit, no trouble should be experienced.

The all-round sing depends on accuracy of the hybrid balance. This, of course, involves careful adjustment of the artificial network which is connected across one side of the hybrid coil to be equivalent to the land line. Generally speaking, singing is more difficult to prevent with high gain receivers, as one might expect.

## Echo Suppressors.

As a general rule, therefore, suppressors are included in the circuit, so arranged that as soon as the local subscriber commences to speak the receiver is cut off the line and speech is routed through to the transmitter. If the distant subscriber speaks, the receiver side of the circuit obtains

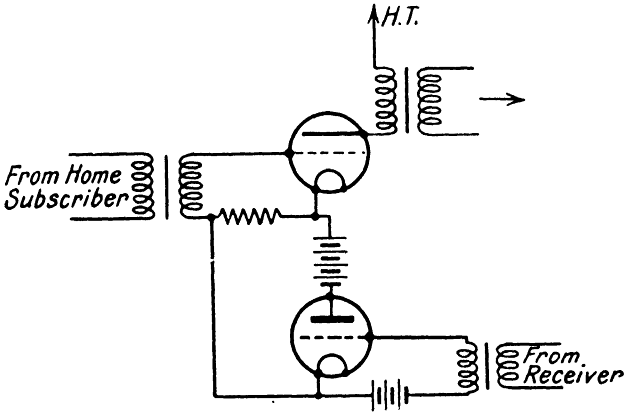



FIG. 114. ECHO SUPPRESSOR CIRCUIT

command, provided the home subscriber is not speaking, cutting off the transmitter again.

This may be done by means of voice-operated relays. The speech frequency is rectified, and the d.c. produced is passed through a relay coil which pulls up an armature and changes over suitable contacts.

Alternatively, electrical methods may be used. Fig. 114 illustrates one such system. In the grid return of one of the amplifying valves is a high resistance, which is in the anode circuit of a valve arranged as an anode-bend rectifier. This normally carries no current, but on the application of speech voltages to the grid, anode current flows and the voltage drop produced by this current in the high resistance sets up a large voltage. This is in such a direction that it increases the bias on the amplifying stage to such an extent as to render the valve inoperative.

## Volume Compression and Expansion.

The power radiated by a single side-band transmitter, and hence the field strength at the receiver, is continually varying in accordance with the modulation. Moreover a large proportion of normal speech sounds is of low intensity, and it is only occasionally that the level rises to full volume, so that the available power is only in use for a small portion of the total time.

An immediate improvement would result if the speech waves could be compressed so that the difference between normal and peak values was reduced. If this compression were in the ratio of 2 : 1 only, the average level would be doubled and the power radiated would increase four times.

Commercial long distance channels, therefore, utilize a device called a *compandor*, which is a portmanteau word denoting a compressor (at the transmitter) and an expandor (at the receiver). Both devices are similar but act in opposite directions.

They comprise an audio-frequency amplifier of which the gain can be controlled by the application of a suitable voltage. For example, two vari-mu valves in push-pull would provide a gain proportional to the grid-bias, the use of two valves being preferable to cancel out any second harmonic distortion which might arise from the non-linear characteristic. A hexode can also be arranged to provide variable l.f. gain by applying speech input to the modulator grid and the control voltage to the signal grid.

Control voltage is obtained from an amplifier followed by a detector, which develops a d.c. voltage proportional to the instantaneous speech level, and by applying this to the non-linear amplifier the modulation may be compressed or expanded (according to the direction in which the control voltage is applied) to any desired extent within reason.

In practice a compression of about 2 : 1 suffices, and at the receiver the modulation is expanded in the same ratio to restore it to normal.

## Frequency, Phase, and Pulse Modulation.

Variation of the amplitude of the carrier is not the only method of transmitting intelligence. Three other methods

are in current use, though their application is usually to specialized systems where some particular advantage is derived and the greater part of the communications of today still use amplitude modulation. They are—

(*a*) *Frequency Modulation.* Here the carrier amplitude is constant but the frequency is varied over a small range on either side of the mean value. The extent of the change is proportional to the depth of modulation required, while the modulation frequency determines the *rate* at which the carrier frequency variations take place.

At the receiving end the frequency variations of the carrier must be converted into amplitude variations, which is done by using a circuit of which the output is proportional to frequency. A normal tuned circuit, operating at a point on the *side* of the resonance curve instead of the peak constitutes such an arrangement and *discriminator circuits*, as they are called, are based on this principle.

The main advantage of frequency modulation is the improvement which it gives in signal/noise ratio because the background noise in a receiver, whether arising from atmospheric disturbances or internally-generated "noise," results from changes in the amplitude of the carrier. But the discriminator is relatively insensitive to amplitude change so that it disregards this interference. Moreover, we can insert limiting devices in the circuit prior to the discriminator which reduce the received carrier to a constant level and very marked improvement in signal/noise ratio can be obtained as a result of this property.

(*b*) *Phase Modulation.* This is similar to frequency modulation but the relative phase of the carrier wave is altered. Alteration of frequency is accompanied by change of phase and the difference between the two systems is mainly that in phase modulation it is the change of phase which is deliberate and controlled by the modulation.

(*c*) *Pulse Modulation.* The development of radiolocation required the generation of very short pulses of carrier. When this technique had become established it became apparent that information could be conveyed by transmitting pulses at controlled rates. Thus the spacing between the pulses, or their duration could be varied in accordance

with the modulation required while still maintaining the pulse amplitude constant and a technique has developed on these lines.

All these specialized forms of modulation are mainly applicable to very high frequencies and are discussed more fully in *Short Wave Radio* (Pitman) by the present author.

# CHAPTER XII

## SHORT-WAVE OPERATION

RADIO communication has been revolutionized by the introduction of short waves. In the early days it was found that radiating properties of a circuit were dependent upon the fourth power of the frequency, and therefore, at first sight, it appeared that the higher the frequency (and the shorter the wavelength) the better the transmission would be. Unfortunately, it was found that short wavelength radiations were very rapidly absorbed. The earth took its toll, and trees, houses, hills, and similar obstructions caused such a rapid deterioration in the ground wave that reception was limited.

The pendulum swung in the other direction and longer and longer wavelengths were used, irrespective of the enormous power required to give the satisfactory field strength, and at one time Bordeaux was using a wavelength of 23 000 metres, corresponding to a frequency of 13 000 cycles, which is actually audible to some ears!

Meanwhile, amateur experimenters had been given the short waves to play with, rather with the feeling that as these waves were of no use for commercial purposes they could be used by amateurs to their hearts' content. And then came the discovery that some freak ranges were being accomplished on these wavelengths. In 1924, Mr. Goyder, a Mill Hill amateur, established communication with Mr. Bell of Dunedin, New Zealand. The possibilities of these wavelengths began to strike the commercial world.

We know to-day that the principal value of these short waves lies in the indirect ray, and not the ground ray.* The waves transmitted into the upper atmosphere are reflected from the Heaviside Layer with much less attenuation than longer wavelengths. At a frequency corresponding to 214 metres the electrons in the upper atmosphere

* The reader should refer to Volume I, Chapter XXIX, for a discussion of wave propagation.

resonate with the magnetic forces of the earth's field, and any transmission at this frequency encounters very heavy absorption of energy. Beyond this point conditions begin to get better again, and at the normal wavelengths now in use, ranging from 15–50 metres, the attenuation of the wave is surprisingly small, the main difficulty arising from the fading which is due to the distortion of the plane of polarization of the wave as already explained in Volume I.

So little is the attenuation, in fact, that signals will sometimes travel two or three times round the world before

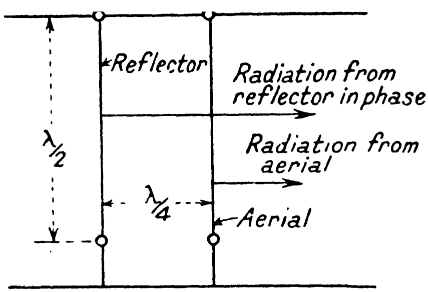



FIG. 115. SIMPLE REFLECTOR ARRANGEMENT

they die out, and this constitutes quite a difficulty in certain instances of high-speed telegraphic working, because there is a time-lag of one-seventh of a second between the original signal and its echo, which may result in a disfigurement of the characters.

**Reflectors.**

Owing to the short wavelengths employed it becomes possible to concentrate the radiation in a beam, and for commercial purposes directional transmission and reception is practically always employed. This is usually accomplished by placing behind the aerial a reflector, which consists of an exactly similar aerial, spaced a quarter of a wavelength to the rear as in Fig. 115.

Currents are induced in this reflector by those in the aerial, and the reflector, therefore, also radiates energy. Due to the space displacement of the reflector, however, the wave radiated by the reflector is out of phase. By

the time the reflected wave has travelled as far as the aerial, it is in phase with the wave being generated there, so that the transmission in the forward direction is strengthened. The wave from the aerial to the reflector, however, which is travelling in the opposite direction, arrives out of phase with the wave generated by the reflector, so that radiation in the backward direction is practically cancelled out.

**Aerial Arrays.**

The simplest form of short-wave aerial is a half-wave arrangement as shown in Fig. 116. This may be fed with

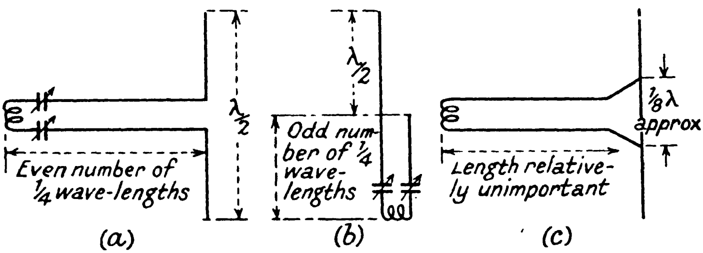



FIG. 116. METHODS OF FEEDING A HALF-WAVE AERIAL

a tuned feeder as at (*a*) or (*b*) (the latter often being known as a "Zeppelin" or "Zepp" aerial), or it may be fed with an untuned feeder, in which latter case the tapping points should be about one-eighth of a wavelength apart. This gives an impedance approximately equal to that of the line, assuming the usual value of 600 ohms. With tuned feeders the lines are usually made a little longer than the correct value, and small tuning or phasing condensers are inserted at the transmitting end and adjusted to give maximum current.

The difficulty about a simple *dipole aerial*, as a half-wave aerial is sometimes called, is that the effective height is limited. One cannot make the length of the aerial more than about 80 per cent of the half wavelength, and even then the current is not uniformly distributed over the wire, so that the equivalent height is again only about 80 per cent of the actual height.

No improvement is obtained by using a longer aerial. Consider a wire $1\frac{1}{2}$ wavelengths long, as shown in Fig. 117. Here radiation would be obtained from the top and bottom portions, and reverse radiation from the middle portion, which would cancel things out and leave the resulting radiation much the same as before. On the other hand, if we could reverse the direction of the current in the centre

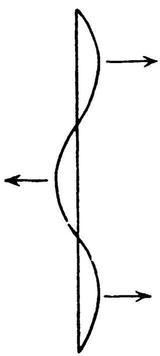


Fig. 117. Illustrating Loss of Radiation with Long Aerial

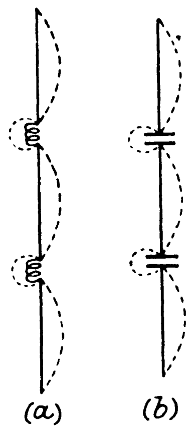



Fig. 118. Tiered Aerials

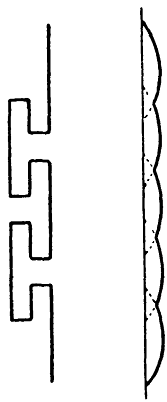


Fig. 119. Franklin Aerial

portion, then we should have all three half-wave aerials adding up to give a combined radiation. This can be accomplished by inserting phasing impedances either in the form of inductances as shown in Fig. 118 (*a*) or capacitances as at 118 (*b*). Such an aerial is known as a *tiered aerial.*

Still better results can be obtained by arranging to bend the wire backwards and forwards on itself, so that the only effective parts are the quarter-wave-length portions in the middle of each half-wavelength section, in which the current is a maximum and fairly uniform. Such an aerial is shown in Fig. 119, and was developed by C. S. Franklin of the Marconi Co.

Now it is possible to carry this procedure out still farther

by arranging more than one aerial and providing each aerial with current in exactly the right phase so that the various radiations add up. This has led to a large number of different aerial formations, usually known as *arrays*, and there are various forms which need not be enumerated here. For further information the reader is referred to *Short Wave Wireless Communication*, by Ladner and Stoner, or to

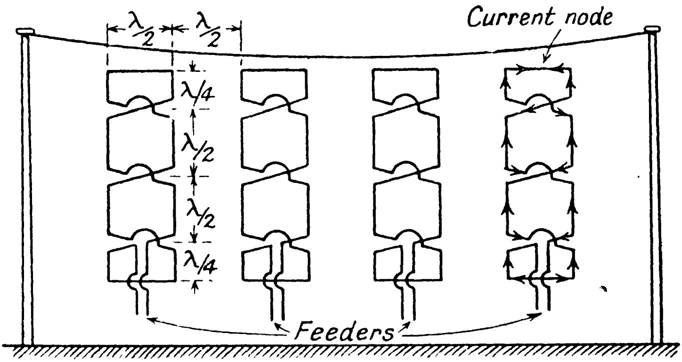



FIG. 120. "STERBA" AERIAL ARRAY FOR SHORT-WAVE TRANSMISSION

*Short Wave Radio* by the present author. Fig. 120 shows one such array. It would be backed by a similar array one-quarter of a wavelength in the rear, forming a reflector.

Fig. 121 is a view of the aerial system at the Grimsby Transmitting Station. There are two aerials with a common reflector system down the centre. They are used alternately to transmit to Australia eastwards during certain periods of the day and westwards at other times, according to conditions. Note the phasing coils in the aerials, the elaborate balance weight system to keep the wires taut even in a wind, and the coupling boxes connecting the tubular (concentric) feeders to the individual aerials. Two aerials are supplied from each coupling box.

## Short-wave Circuits.

We now come to a consideration of the transmitting circuits themselves. These are, in general, similar to those

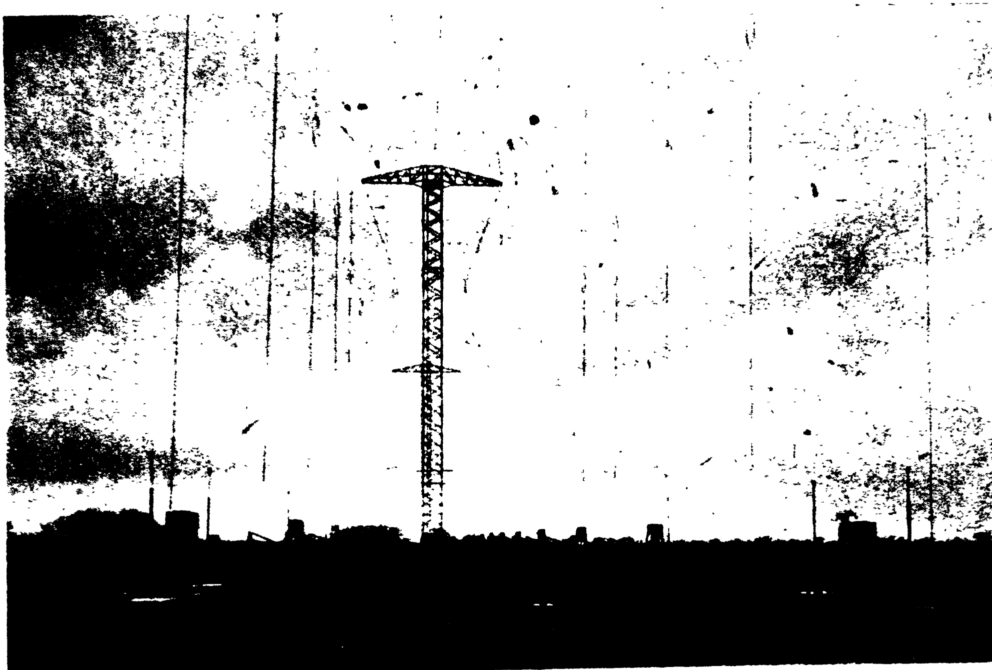


FIG. 121. VIEW OF THE AERIAL SYSTEM AT THE GRIMSBY SHORT-WAVE STATION
*(By courtesy of Marconi's W. T. Co.)*

outlined in Chapter I, with special precautions to allow for the high frequencies in use. The coils, of course, are small, and are usually made of copper tube, since the very high frequency current only flows in the skin of the wire.

Alternatively, strip conductor may be employed, wound with the flat surface parallel to the axis of the coil, thus approximating to a current sheet. The conductors should be liberally designed and closed loops in the wiring or chassis avoided since they might absorb energy from the oscillating circuit. Wave changing cannot be carried out by short-circuiting the end turns as is usual with lower frequencies, it being necessary to change over completely from one coil to another or alternatively to employ a series–parallel arrangement.

Insulators must be used sparingly and kept, where possible, outside the high-frequency field. The coils are thus always air-spaced and mounted on skeleton formers while the general structure must be self-supporting as far as is practicable, requiring a minimum of mechanical fixing. Sulphur-filled fittings are to be avoided, as the sulphur heats up under the influence of the dielectric currents. Ebonite contains an appreciable portion of sulphur and is thus unsuitable and porcelain or similar ceramic material is the most satisfactory.

### Valve Design.

Valve capacitances are inclined to be troublesome, since the total tuning capacitance cannot be large or no voltage would be developed across the circuit, again owing to the high frequency. Hence, the valve capacitance itself constitutes a considerable proportion of the total capacitance, and this, in turn, means that the valve electrodes are carrying a considerable proportion of the oscillating current (perhaps more than 50 per cent).

This led to trouble in the early days, because the wires leading out through the glass of the valve were not made thick enough, and they used to overheat and melt the glass. Specially thick leading-in wires are now used in short-wave valves.

A further trouble was the development of hot spots on

the glass bulb, due to eddy currents set up in small metallic deposits formed on the glass during the evacuation. This led to puncture of the glass and consequent failure. One remedy was to enclose the valve in a gauze sheath since this produced uniform eddy currents over the whole surface with consequent distribution of the heat, but improved

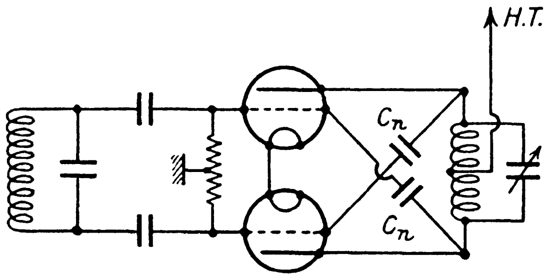



FIG. 122. FRANKLIN BRIDGE CIRCUIT

technique has largely overcome the difficulty and such special precautions are now not necessary.

**Neutralizing.**

Very symmetrical neutralizing circuits have to be used in the amplifying stages, even the leads themselves being maintained exactly the same length. Otherwise, the phase of the voltage fed back by the neutralizing circuit is not exactly opposite to that fed back through the valves, and incomplete stability is obtained. A modification of the Rice circuit (Fig. 32) is usually adopted.

C. S. Franklin has developed a special bridge circuit using two valves in push-pull as shown in Fig. 122. This circuit is an absolute balance both as regards voltage and phase, even down to quite short wavelengths. Even here the length of the leads may be sufficient to introduce enough inductance to throw the phase balance out, in which case condensers have to be inserted in the offending lead, shunted, if necessary, by a high-frequency choke to provide a d.c. path.

Within the last few years special screen-grid valves

have been developed for this short-wave technique. These valves are designed for class B operation (positive drive) in the usual way, and are also provided with special lead-in wires to avoid any overheating.

## Parasitic Oscillation.

A transmitter will sometimes be found to be oscillating in more than one mode, producing what is called a parasitic oscillation. This may be small or large, depending on the conditions of operation and in severe cases may swamp the true oscillation altogether.

Such oscillations are evidenced by abnormal conditions in the circuit such as high anode feed currents, excessive grid currents—causing the grid to become red hot, instability, etc. It may be found that the circuit shows all the symptoms of strong oscillation, but the actual oscillation frequency cannot be found within the normal range by the wavemeter or other measuring device.

Such parasitic oscillations are usually found in multi-valve transmitters, though they may occur with quite simple circuits, especially when using valves of high mutual conductance. They are of two general types.

(*a*) Oscillations near the fundamental frequency. These are usually due to bad layout. The leads between the valve and the tuned circuit may form an oscillating circuit of lower resistance than the correct circuit (particularly on ultra-short waves) and the valve will always choose to maintain oscillation in the easiest mode. Alternatively, a lead too near the framework may form an oscillating circuit with its capacitance to earth. The remedy is to check the layout carefully for such possibilities.

(*b*) Very high frequency oscillations. These are due to self-oscillation in the leads to the valves. They may occur in the transmitting, amplifying, or even modulating circuits, and in the latter case may generate considerable power.

On longer waves it is possible to check such parasites by including resistances of a few hundred ohms in the anode and/or grid leads but this is not practicable on short-waves, and the only remedy is to pay special attention to layout, keeping the leads short, and avoiding leads of

similar length in adjacent stages so that the natural frequency of such circuits may be different throughout the chain.

## Frequency Multiplication.

The importance of maintaining an absolutely steady frequency has already been mentioned in Chapter I. It is more than ever important on short waves. A telegraphic communication channel may occupy perhaps 100 cycles.

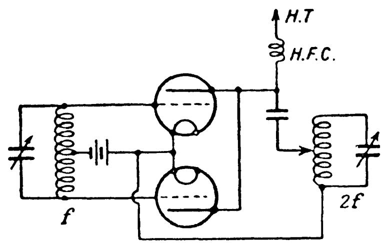



Fig. 123. Frequency Doubling Circuit

A drift of 0·1 per cent at 10 megacycles (30 metres) is 10 000 cycles, one hundred times as much as the frequency band. The constant-frequency circuits quoted in Chapter I are only practicable up to frequencies of two or three megacycles, beyond which point their application becomes increasingly difficult, and it is customary therefore to generate oscillations in the first place at a lower frequency than is required, and to multiply the frequency subsequently. An over-biased push-pull circuit will produce the necessary doubling action, the anode circuit being tuned to twice the input frequency.

Such a circuit is shown in Fig. 123. The anode circuits are arranged in phase instead of in opposition, as usual, so that each valve acts as a half-wave rectifier. It is, of course, possible to double the frequency with only one valve (over-biased to give the rectification necessary) but this is only half as efficient.

The arrangement of Fig. 123, however, is unsymmetrical, and that of Fig. 124 is to be preferred for short-wave work.

Here a centre-tapped anode circuit is used and the arrangement is fully symmetrical. As it stands it will not double the frequency, but if one of the cathodes is extinguished, and the anode circuit is tuned to twice the frequency of the input, oscillations are maintained satisfactorily if the input

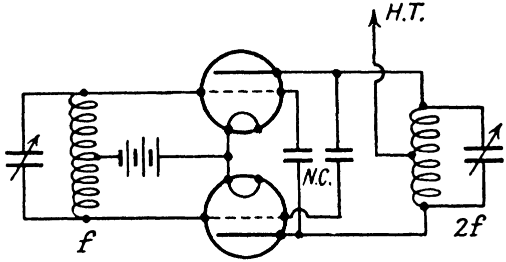



FIG. 124. NEUTRALIZED FREQUENCY DOUBLER

is overbiased. It is even possible to tune the anode circuit to three times the input frequency and so obtain a trebling action.

Neutralizing is still desirable for the circuit may oscillate of its own accord at the double frequency which is not desirable. The circuit of Fig. 124 shows a neutralizing arrangement obtained by feeding energy from the anode of one valve back on to the grid of the other valve, which provides the necessary phase reversal.

## Crystal Control.

Reference should be made in conclusion to the use of quartz crystal control for high-frequency transmitters. It is found that a quartz crystal is found to exhibit what s known as a *piezo-electric effect.* If the crystal is subjected to strain, either tension or compression, electric charges will appear at the edges. Conversely, an alternating potential applied across the crystal will cause it to expand and contract. Furthermore, there is a particular frequency at which this expansion and contraction take place most readily, so that a crystal of this nature acts as a mechanical oscillating circuit.

The resonant frequency is high, ranging from several hundred thousand cycles up to something of the order of megacycles, depending upon the dimensions of the crystal. Such a crystal between two plates forms a mechanical radio-frequency oscillating circuit which can be maintained

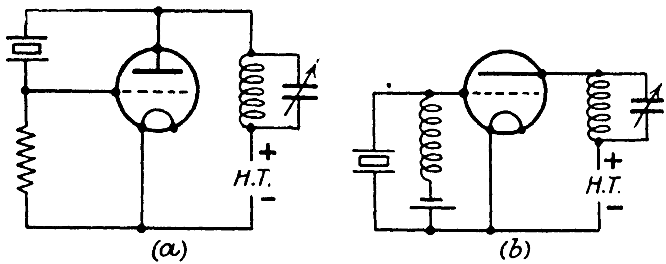



FIG. 125. CRYSTAL-CONTROLLED OSCILLATOR CIRCUITS

by a valve, and which will have a very nearly constant frequency provided the dimensions of the crystal are maintained the same. This is done by keeping the atmospheric conditions, particularly the temperature, as constant as possible, the whole crystal being often enclosed in a special chamber. A frequency stability of ·01 per cent is found quite feasible.

Fig. 125(*a*) shows a circuit for a crystal controlled oscillator. The anode circuit is tuned so that it shall present a reasonably high impedance to that of the frequency under consideration, but the actual oscillating frequency is controlled entirely by the quartz crystal across anode and grid. An alternative method is to connect the crystal between grid and filament, biasing the valve through a h.f. choke as shown in Fig. 125(*b*). In both cases, the frequency of the tuned circuit must not be too near that of the crystal, or "pulling" may take place, causing instability, but if this precaution is observed no difficulty is experienced.

The crystal slice has to be cut in a particular way, the length which determines the frequency being measured along a particular axis. Quartz crystals are hexagonal in cross section. If we call the axis joining two opposite corners the $X$-axis and the axis at right angles to this

the $Y$-axis, we find that the best results are obtained with slices cut from the crystal with sides parallel to the $Y$-axis. This is known as a $X$-cut crystal since its face is normal to the $X$ axis.

Such a crystal will have two modes of oscillation, one corresponding to vibration along its length $y$ in the $Y$-axis direction and the other (higher) oscillation corresponding to its breadth $x$ in the $X$-axis direction. There is no oscillation corresponding to the $Z$-axis.

The frequencies of these oscillations are given by

$$f_y = 2730/y \text{ and } f_x = 2730/x$$

where $f$ is in k/c, and $x$ and $y$ are in mm.

A crystal cut with its sides parallel to the $X$-axis has only one mode of oscillation—that corresponding to its depth in the $Y$ direction and it is known as a $Y$-cut crystal in consequence. This is given by

$$f = 2070/y$$

where $y$ is in mm. as before.

$X$-cut crystals have a negative temperature coefficient while $Y$-cut crystals are positive. Recently a cut known as the *AT* or constant temperature cut has been evolved in which these two opposite effects are caused to offset one another giving a crystal which does not require temperature correction.

There are also numerous forms of circuit arrangement some of them arranged to generate frequencies which are a multiple of the fundamental crystal frequency, this being known as *harmonic excitation*.

For further information on points of detail the reader should refer to *Radio Engineering* by Terman or *Short Wave Radio* by the present author.

## Short-wave Receiving Aerials.

Though short-wave transmissions can be received on a normal fairly short aerial, it is preferable to use a special arrangement, and some such device is always used for commercial reception. For beam reception an array would be used similar to those described for transmission, but

for broadcast reception a simpler aerial is used, usually more or less directional in character.

The Beverage aerial described in Chapter VIII is particularly suitable for short waves. The length of such an aerial should be several wavelengths, which runs into miles at medium or long wavelengths but is only a few hundred

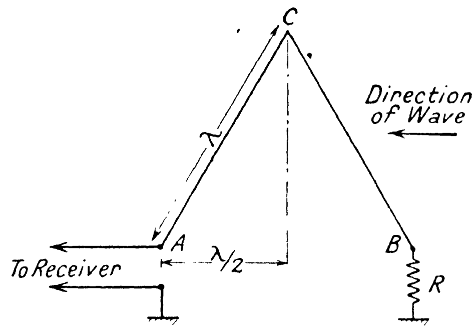



FIG. 126. V AERIAL FOR SHORT-WAVE RECEPTION

feet on short waves. The forward end is earthed through a resistance equal to the surge impedance of the line, the receiver being coupled to the far end through a matched transformer.

The directivity of such an aerial is not dependent on the wavelength which enables it to be used over a wide range of frequencies. Two such aerials side by side are often used, while several may be employed separated by several wavelengths for diversity reception.

## Diamond Aerial.

Another form of aerial capable of sharp discrimination between forward and backward reception is shown in Fig. 126. Consider a vertical wire acted upon by a wave, and terminated at the earth end by its characteristic impedance. The voltages induced in the different parts of the wire will all be slightly different in phase on arrival at the termination because of the increasing distance which they have to travel, but they will nevertheless add up to

give a positive resultant provided the wire is less than $\lambda/2$ long.

If the wire is exactly $\lambda/2$ in length the resultant will be a maximum. Beyond this point the currents coming from the far end will be out of phase, causing the resultant to decrease again and if the wire is one wavelength long the resultant becomes zero. If, however, the top of the

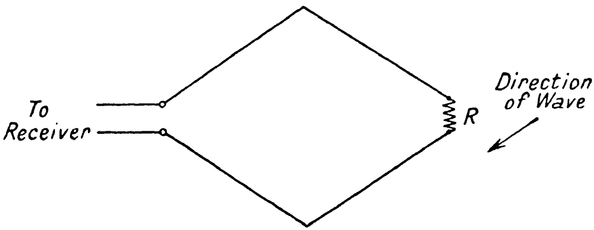



FIG. 127. HORIZONTAL DIAMOND AERIAL

wire is advanced half a wavelength into the wave direction, the currents at the top will receive half a wavelength start and will therefore arrive at the bottom in phase and so for all the elements down the wire, so that maximum current will again be obtained.

The whole aerial comprises two such wires, tilted at the correct angle. A forward wave then builds up continuously and gives maximum response, but a wave from the wrong direction or of different wavelength does not build up. Moreover, waves from the back direction build up their voltage at the point $B$, where they are absorbed by the resistance $R$ without reflection. The successful operation of this discrimination, however, involves careful adjustment of $R$.

## Horizontal Diamond.

A modification of this is the horizontal diamond shown in Fig. 127. Here a symmetrical arrangement is employed which avoids earthing and makes the adjustment of $R$ simpler. The sides are made several wavelengths long which nullifies the wavelength selectivity over a wide band, but retains the directional discrimination. The actual

length of side and the angle of the diamond depends on the inclination to the ground of the wave-front to be received. It is, of course, only responsive to the horizontal component of the received wave of which there is usually an ample supply, owing to twisting of the plane of polarization during reflection from the Heaviside Layer.

There are numerous other forms of aerial, but it is not considered desirable to discuss them at length. The reader who desires further information should refer to *Short Wave Wireless Communication*, by Ladner and Stoner, "Beam Arrays and Transmission Lines," by Walmsley, *Journal I.E.E.*, Vol. 69, Feb., 1931, and "Developments in Short-wave Directive Antennae," by Bruce, *Proc. I.R.E.*, Vol. 19, August, 1931.

**Short-wave Receivers.**

Short-wave receivers follow ordinary receiver technique with due allowance for the much higher frequency. The coils used should be wound with stout wire and the insulating material chosen with care to minimize losses.

A simple receiver would comprise an h.f. isolating stage feeding a reacting detector followed by l.f. stages as required. Simple detector-l.f. sets are of limited use because of the greater effect of aerial variations at short waves. Any variation of capacitance, due to swaying in the wind or similar causes, will alter the frequency by a sufficient amount to produce large variations in the heterodyne note (assuming c.w. reception) or of the tune itself which would adversely affect a critically adjusted telephony receiver. A buffer h.f. stage minimizes both these defects.

The h.f. gain obtainable is limited, mainly due to the low dynamic impedance of the coils. The inductance is only a few microhenries so that $L/CR$ is rarely more than 10 000 ohms, giving a stage gain with a screened valve of the order of 20 and often less. This, however, is quite sufficient for the purpose. A coupled aerial winding will give an additional step-up of 2 or 3 if properly designed, though with imperfect design it is easy to produce a step-down rather than a step-up.

More sensitive receivers may employ several h.f. stages,

each giving a small gain of 10 to 20. Such arrangements are practical particularly for reception over limited bands of frequency. The high frequency of the currents involved, however, requires particularly good decoupling condensers. The usual paper condensers are not always satisfactory and mica condensers of $0{\cdot}01\,\mu$F. capacitance are preferable.

The more usual type of receiver for long-range reception, however, is the superheterodyne using an i.f. of 400–500 k/c. If two or more signal frequency circuits are employed prior to the frequency changer (preferably in association with an h.f. stage) the signal-frequency selectivity is sufficient to avoid second channel interference and the necessary adjacent channel selectivity is, of course, obtained from the i.f. stages. Efficient a.v.c. is a considerable advantage owing to the serious fading which is inseparable from short-wave reception which is almost entirely carried out by the "sky" wave.

The characteristics of short-wave transmissions and the mechanism of reflection from the Heaviside and Appleton layers are discussed in Volume I, Chapters XXIX and XXX.

# CHAPTER XIII

## ULTRA-SHORT WAVES

THE increasing use of higher and higher frequencies has resulted in the technique of transmission outstripping that of reception by an appreciable margin. The usefulness of short waves for long-distance communication ceases at about 30 megacycles (10 metres), because beyond this point the waves are insufficiently refracted (i.e. bent round) at the Heaviside Layer to return to earth again. Wavelengths below this limit, therefore, are usually termed *ultra-short* waves, and their range is restricted to optical limits. In other words, there is little bending round the surface of the earth, and any serious obstruction in the path of the rays will act as a complete barrier.

Prior to 1939 the main use of ultra-short waves was for limited range point-to-point and television transmissions. The need for ever-shorter wavelengths for radiolocation purposes however, stimulated the development of transmissions on wavelengths of 10 cm and later on 3 cm with the result that a great deal was learned of the generation, reception and propagation of such waves.

As far as propagation is concerned, it has been established that for metre wavelengths, i.e. from 10 m down to just under 1 m, the transmission is substantially optical and reception is limited to a range not much beyond the visual horizon. At centimetre wavelengths, however, a phenomenon known as *super-refraction* is often experienced. This arises under conditions of *temperature inversion*.

Normally the air temperature decreases approximately 1°C for every 300 ft. rise in elevation. Sometimes, however, the temperature may rise as the altitude increases, at any rate for the first few hundred feet. This provides a layer of warm air on top of the cool air and this layer, together with the surface of the earth, constitute a duct or guide

which will conduct the waves for distances far in excess of the visual horizon.

Detailed information on this subject was given at the I.E.E. Radiolocation Convention in March 1946, notably by Booker (*Journal I.E.E.*, 1946, **93**, Part IIIA, p. 69 *et sqq.*)

The limited range of the waves is in some respects an advantage, because one of the principal difficulties to-day, particularly with broadcast transmissions, is the interference caused by stations outside the area which they are intended to serve. The purely local character of ultra-short wave transmissions has been put forward as a solution to the broadcasting problem, while there are numerous other applications in which the local character of the radiation is helpful.

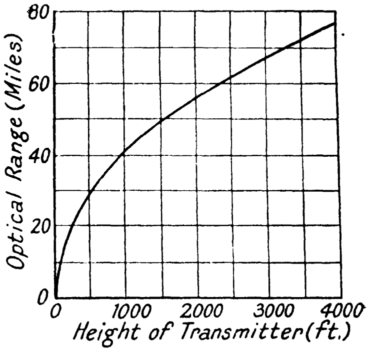



FIG. 128. VARIATION OF OPTICAL RANGE WITH HEIGHT

The range of the transmission, calculated on an optical basis, in terms of the height of the transmitter is shown in Fig. 128.

A further advantage of ultra-short wavelengths is that, owing to the high carrier frequency, wide modulation bands can be used, and this is particularly helpful for television and similar purposes where a band of several hundred kilocycles may be required. On the other hand, the undoubted effect of any obstruction renders reception more difficult. Moving the aerial from the side to the top of the building may make all the difference between successful reception and the reverse.

## Generating Circuits.

The generation of ultra-short waves can be carried out by ordinary oscillating valves down to limits of about 1 metre. The coils have to be quite small and the layout

of the wiring has to be carefully considered, as the inductance of any loops in the wiring may exceed that of the coils themselves. In fact, it is often better to arrange the wiring in the form of a fairly small loop and to use this as the inductance.

Power amplification at these frequencies is difficult, but drive circuits are still desirable from the point of view of frequency stability, and also to obtain as much power in the aerial as possible owing to the limitations of the self-oscillating coupled circuit.

## Reception of Ultra-short Waves.

The design of receivers for ultra-short waves is a matter of some difficulty, though the technique is developing fast. The simple reacting detector is of use at short ranges if a rigid aerial is employed, though the super-regenerative detector is more sensitive. This arrangement, described in more detail in Volume I, comprises a detector which is made to oscillate but is quenched at a high audio or supersonic frequency by suitable means. The oscillation then builds up to a value proportional to the signal picked up and will thus follow the modulation of the received signal satisfactorily and will give large amplification.

The essence of successful operation is that the quenching must occur before the oscillation has reached its limiting value. Otherwise the final value before quenching will clearly not be proportional to the modulation of the received signal. The selectivity of the system is poor and it is apt to be noisy owing to the large gain which brings up background noise equally with the signal.

Straight h.f. amplification has been achieved at these wavelengths by using small capacitances to obtain a reasonable $L/C$ ratio. This involves the use of special valves having no caps or holders of the usual type, the leads being brought out through the sides and taken direct to the circuit via skeleton holders which introduce no appreciable length of lead and have negligible capacitance. The valve structure is also on a small scale since this reduces the valve capacitance considerably. The valve itself looks

like an acorn and is about the same size, so that such valves have been called "acorn valves."

Alternatively the superheterodyne principle may be used, the frequency being converted to 10 or 20 Mc/s at which stage gains of 30 to 50 are possible. This gain may have to be limited if a wide band-width is required, as explained in Chapter XV on Television.

Special half-wave aerial systems are desirable, located at a height of one or more wavelengths to bring them above the level of local screening and to avoid, as far as possible, the interference from motor-cars, the ignition systems of which radiate damped waves on about 7 metres. The aerial is coupled to the set through suitable feeders of the type described in Chapter III.

### Electronic Oscillations.

Below about 1 metre the frequency becomes so high that it is comparable with the speed at which the electrons are moving inside the valve. This being the case, there is an appreciable time lag in the response of the electrons, and the phase relations are no longer satisfactory for the maintenance of the oscillation. An entirely new technique is required for the production of these *micro waves*, as they are sometimes called.

One method is by use of the Barkhausen-Kurz effect, which was discovered by two German engineers during experiments on the vacuum of transmitting valves. If the grid of the valve is at a high positive potential, electrons are attracted across from the filament. They will shoot through the grid, but are prevented from reaching the anode by the fact that the voltage on that electrode is less than that on the grid, being actually zero in some cases. They therefore return to the grid, but again shoot through the spaces until the positive voltage on the grid manages to pull them up and they commence to return once again.

Thus, the electrons inside the valve are themselves oscillating backwards and forwards around the grid, and they do this at a very high frequency, dependent upon the dimensions of the valve itself.

It might appear that the various electrons would all oscillate at random, the total effect being little or nothing due to cancellation. Actually, however, a random distribution is unstable and any deviation from the normal motion in unison sets up forces which rapidly bring the offending electron to rest.

A good treatment of the subject will be found in a paper by Hollman entitled "On the Mechanism of Electron Oscillations in a Triode," *Proc. I.R.E.*, February, 1929.

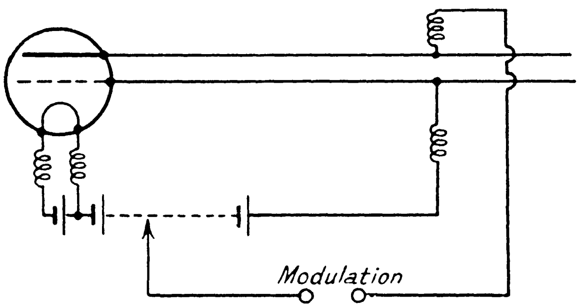



FIG. 129. BARKHAUSEN-KURZ CIRCUIT

**External Circuit.**

The appearance of current in the external circuit arises from the movement of the electrons past the grid. The electrons in the grid wires are alternately repelled and attracted as the electron cloud in the valve surges back and forth.

An external circuit consisting of a Lecher wire is connected to grid and anode, the direct voltage supply to the electrodes being fed in at a voltage node through suitable radio-frequency chokes. The circuit is shown in Fig. 129. The influence of the external circuit on the frequency of the oscillation is not altogether agreed by various investigators, but it has, at any rate, to be tuned approximately to the frequency determined by the dimensions of the valve and the voltages on the electrons. Very roughly, the wavelength is given by the expression

$$\lambda_{cm} = \frac{2\,000}{\sqrt{V_g}} \cdot \frac{AV_g - BV_a}{V_g - V_a}$$

where $V_g$ and $V_a$ are the grid and anode voltages, $A$ is the distance from anode to filament and $B$ is the distance from grid to filament.

## Gill-Morrell Oscillations.

A slightly different form of oscillation, also dependent upon the valve, was discovered by Gill and Morrell more

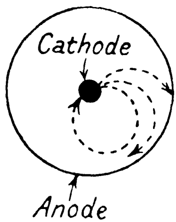



FIG. 130. ILLUSTRATING ACTION OF MAGNETRON

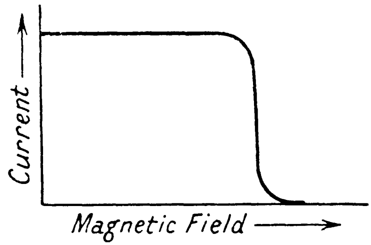



FIG. 131. MAGNETRON CHARACTERISTIC

recently. Here the anode of the valve was considerably farther away than usual, and the oscillation obtained did not agree with the Barkhausen-Kurz theory. The discoverers explain this oscillation by assuming a sort of compression and rarefaction of the space charge inside the valve, but the fact remains that, as before, the oscillations are controlled by the voltages on the electrodes and are led out into a radiating system through a suitable tuned feeder. For further information see Megaw, *Journal I.E.E.*, Volume 72, April, 1933.

## The Magnetron.

Another form of ultra-short wave oscillator is the *magnetron*, in which the movement of the electrons is controlled magnetically instead of electrostatically. This type of valve was actually developed as far back as 1920 by Dr. Hull in America, but it was not a satisfactory proposition at ordinary radio frequencies. The valve consists of a cathode and anode, while an axial magnetic field is produced

by suitable external means. The arrangement is illustrated in Fig. 130.

The operation is as follows. The electrons leave the cathode and proceed to travel straight to the anode. The axial magnetic field, however, deflects them at right angles, as shown in Fig. 130. If the field is strong enough, the electrons never reach the anode at all but travel round in circles. There is thus a very critical cut-off point depending upon

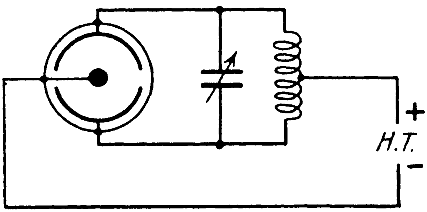



FIG. 132. MAGNETRON CIRCUIT

the strength of the magnetic field, a typical magnetron characteristic being as shown in Fig. 131.

It is clear that the electrons which fail to reach the anode will return to the cathode with some definite oscillation period, and oscillations can therefore be generated in this valve rather similar to the Barkhausen-Kurz type.

If the anode is split into two segments, as in Fig. 132, it will be clear that, under suitable conditions of magnetic field strength and anode voltage, electrons which commence their spiral journey under the influence of an increasing positive field on, say, the bottom half of the anode, actually arrive at the top segment on which the potential is decreasing. In other words a negative-resistance effect is produced whereby a decreasing voltage is accompanied by an increase in current and vice-versa. Under such conditions the external circuit will extract power from the valve and oscillation can be maintained.

Modern valves use anodes having 8, 12 or even 16 segments and pulse outputs of several kilowatts can be achieved at wavelengths as low as 3 cm.*

* See *Journal I.E.E.*, 1946, **93**, Part IIIA, pp. 928, 977 and 985.

**Velocity-Modulation Valves.**

An entirely new technique has arisen in the war period by the use of valves in which electron stream is alternately accelerated and decelerated. Such valves are called velocity-modulation valves, a typical example being the klystron illustrated in Fig. 133.

An electron gun, comprising a cathode and an accelerator, generates an electron stream in a manner similar to that

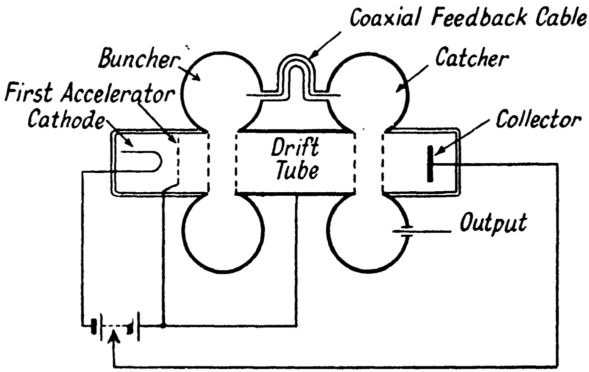



FIG. 133. KLYSTRON OSCILLATOR

in the cathode-ray tube (see Chapter XIV). This stream then passes through a doughnut-shaped metal chamber known as a buncher (shown in section in Fig. 133). This chamber is merely a tuned circuit in which the inductance and capacitance are "built-in" since the wavelength is so short that conventional coils and plates have too large an inductance and capacitance.

If this buncher is in an oscillating condition an electric field appears across the central portion, indicated by the dotted lines, which alternately accelerates and decelerates the electrons in the stream. A little farther along the tube is a second resonant chamber known as a catcher, but because of the time taken by the electrons to travel down the drift tube the variations in the velocity of the stream are 180° out of phase with the field across the catcher. Thus again we have a negative-resistance action and the

electron stream delivers energy to the catcher and so maintains the whole system in oscillation.

A small part of the energy generated in the catcher is fed back to the buncher to maintain it in oscillation. The remainder is available for external use.

## Reflectors. Wave Guides.

An important advantage of the use of microwaves arises from their extremely short wayelength. The devices used for reflecting, concentrating and general handling of light waves nearly all operate by virtue of the fact that their physical dimensions are far greater than the wavelength of light.

With microwaves, of the order of 3 to 10 cm, it is clearly feasible to construct apparatus of which the physical dimensions are at least of the same order as the wavelength concerned and this permits the construction of parabolic reflectors, reflecting curtains and similar arrangements by means of which the waves may be handled quasi-optically.

In particular it is possible to use hollow metal tubes to carry the waves. Such tubes are known as *wave guides* and they provide a means of transmitting energy at frequencies where the usual form of transmission line begins to prove inadequate. At 10 cm, a coaxial cable or a wave guide are of the same order of efficiency though wave-guide technique has certain advantages which often make it preferable. At 3 cm the wave-guide is definitely superior and is almost invariably used.

Wave guides may be circular or rectangular in cross section. The rectangular form is easier to make and handle, particularly when corners are required, so that this form of construction is most usual. The mechanism by which the propagation takes place along the guide is explained more fully in the author's companion volume, *Short Wave Radio*, (Pitman) which also gives further details on the technique of the generation and reception of ultra-short waves*.

* See also "Transmission of Electromagnetic Waves in Hollow Tubes of Metal" Barrow. *Proc. I.R.E.*, 1936, **24**, p. 1298, and "Wave Guides in Electrical Communications" Kemp. *Journal I.E.E.*, 1943, **90** Part III, p. 90.

# CHAPTER XIV

## MEASUREMENTS IN RADIO COMMUNICATION

THE field of measurement in radio communication ranges from the determination of simple constants such as resistance, inductance and capacitance, and fundamental quantities such as current, voltage and power, up to tests on complete receivers for sensitivity, fidelity, selectivity.

The simpler forms of measurement have already been discussed in Volume I, but we shall consider here some of the more advanced measurements, although a complete discussion of measurement technique is beyond the scope of this work. Inductance, capacitance and h.f. resistance will not be mentioned, as they have already been dealt with.

### Current and Voltage.

Some reference may be made, however, to the methods adopted for measuring current and voltage in communication engineering. For d.c. measurements a moving-coil meter is used, and the only point to be noted here is that the connexion of the meter across the circuit should not affect the reading in question.

For instance, consider the circuit of Fig. 134, where we require to measure the voltage on the anode of the valve. The current taken by the voltmeter will have to flow through the anode resistance, and will therefore increase the voltage drop so that the voltage at the anode will be less than its normal amount. If the voltmeter is an insensitive one it may take more than the valve itself, giving misleading results.

The remedy is to use a sensitive voltmeter which takes only about 1 mA. for full-scale deflection, but even so it is better to measure the voltage across the resistance. If the voltmeter is reading 100 volts maximum and takes 1 mA. for full-scale deflection, it must obviously have an internal resistance of 100 000 ohms. The connexion of such a resistance across an anode resistance of, say, 20 000 ohms

would produce a negligible error (although some error would exist and could be calculated if necessary). Knowing the h.t. voltage, the anode voltage can be deduced by simple subtraction.

This somewhat lengthy example is given to show the importance of allowing for the current taken by the meter

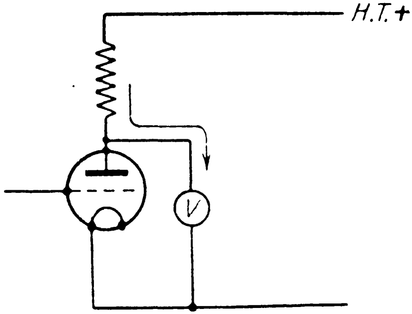



FIG. 134. ILLUSTRATING INFLUENCE OF METER ON CIRCUIT
The current taken by the voltmeter $V$ may vitiate the voltage reading

(or the voltage dropped across the meter in the case of a milliammeter which is inserted in series with the circuit).

## A.C. Measurements.

With a moving-coil instrument the deflection is dependent upon the direction of the current, so that with an alternating current no reading is obtained. For power-frequency measurements moving-iron instruments are often used. These are arrangements of lightly-pivoted armatures which are attracted by a coil to an extent which depends on the current through the coil. By a suitable design a fairly open scale can be obtained. The magnetic pull is proportional to the square of the current, so that the meter is insensitive to small currents but becomes much more sensitive as the current increases. By suitable design, however, the deflection for the larger currents can be limited, so that a fairly uniform scale is obtained.

The moving-iron instrument, however, is not a sensitive

one, it being difficult to produce a movement giving full-scale deflection on as little as 5 mA. Added to this the instrument is only suitable for low frequencies, because the self-capacitance of the winding will introduce resonant effects at higher frequencies.

## Rectifier Meters.

For audio-frequency voltages a different form of meter has to be used. One quite commonly employed is the *rectifier meter*. This consists of a moving-coil meter, the current to which is supplied through a small bridge rectifier

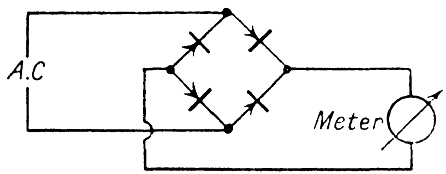



FIG. 135. CIRCUIT OF A.C. RECTIFIER METER

as shown in Fig. 135. Research into rectifier design has evolved types which are free from serious frequency error up to nearly 100 000 cycles per sec., and since a moving-coil meter can be used this form of instrument is quite sensitive.

Its principal disadvantage is that it measures the mean rectified current. What we require to know is usually the r.m.s. current. The instrument can be calibrated on a sine wave to read the r.m.s. value, but this calibration will obviously not hold if the wave form is seriously distorted. In fact, quite a small amount of distortion is sufficient to introduce error of 5 per cent or 10 per cent, and as distortion is quite common in audio-frequency measurements, particularly during investigation, this inherent disability of the rectifier meter should always be borne in mind.

## Thermal Meters.

The next type of instrument is the *thermal meter*. For measurement of large currents of the order of amperes,

a hot-wire meter can be employed. This is a meter in which the increase in length of a thin stretched wire is caused to operate a pointer through a series of silk cords. For more sensitive work, however, the thermo-couple type of meter is used. Here the current is passed through or close to the junction of two dissimilar metals, and the heat generated develops a voltage at the junction. This voltage is applied to a sensitive moving-coil meter which can be calibrated directly in terms of the alternating current flowing through the heater wire.

The thermo-couple may be built in to the instrument or may be separate, the latter course being adopted for larger currents, and also sometimes for smaller ones owing to the ease of replacement. The disadvantage of this kind of meter is that it only has a small overload factor, so that if the current exceeds the full load current of the meter by more than 50 per cent, there is a danger that the heater wire will become so hot that it will burn out and have to be replaced. A special form of bridge circuit has been developed by the Weston Electrical Instrument Co., which will withstand a heavier overload, and this is used in many of their meters.

The thermal meter obeys a strict square law, so that it is insensitive towards the bottom of the scale but gives a wide scale at the top. This is a disadvantage which cannot be avoided. For the measurement of still smaller currents *vacuo-junctions* are used, these consisting of thermo-couples mounted in a glass bulb from which the air is exhausted, as this is found to improve the sensitivity.

Junctions are made giving a full-scale deflection with a current of 1 mA. only. With such junctions, however, the heater wire has to be so fine that the resistance is several hundred ohms, which often has a serious effect on the circuit in which the meter is inserted. Generally speaking, a thermo-junction has a resistance of something under 5 ohms for ordinary work.

Voltages are measurable, of course, by inserting suitable series resistances, but since, even under the best conditions, the current required for full-scale deflection is 10 or 20 mA. the thermal voltmeter is little used.

## Dynamometer Instruments.

For certain classes of work, *dynamometer* instruments are used. These employ an arrangement of two coils, one fixed and one moving. If current is passed through the whole arrangement, the magnetic fields produced interact and set up a force between the coils, causing the moving one to rotate. This movement is indicated by a pointer, and is proportional to the square of the current in the circuit.

Again owing to self-capacitance troubles, these instruments are usually suitable only for relatively low frequencies.

## Power Measurement.

The dynamometer instrument is particularly useful for measuring power. Here the current is passed through one set of coils, usually the fixed one, since this can the more easily carry a larger current, while the other set is connected in series with a high resistance across the source of supply so that it measures the voltage. The resultant force is, therefore, proportional to the product of the current and the voltage, with due regard to their phase relationship, which is exactly what we require for measuring the power.

This class of instrument can, with proper precautions, be used at audio frequencies, but it is not much employed here because the powers to be dealt with are only a few watts, and it is an expensive matter to make a dynamometer wattmeter of a low reading type. The more usual practice for audio frequencies is to measure the current flowing through a known resistance in the circuit, or alternatively the voltage across it, when the power can be calculated from the usual relation, $W = I^2R = E^2/R$.

## Valve Voltmeters.

An instrument much used for voltage measurements in communication work is the *valve voltmeter*, which is simply a rectifying valve with a meter in the anode circuit. Various forms of the instrument are made, of differing sensitivities, depending upon the requirements. Fig. 136 shows a typical "reflex" circuit in which an anode-bend arrangement is used, the bias being obtained by the voltage drop across a resistance in the anode circuit. Consequently, as the

anode current increases due to the application of a signal, the grid bias also increases and tends to restrict the rise of current. In this way, the deflection can be maintained more or less proportional to the input signal. It will be

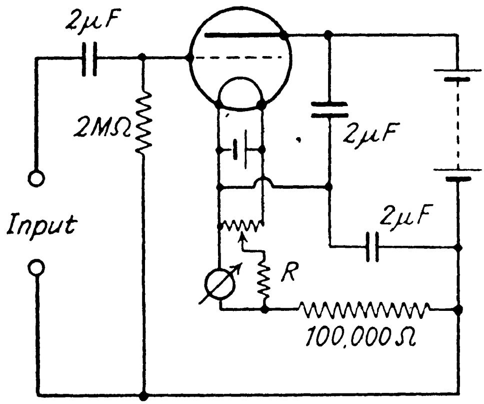



FIG. 136. REFLEX VALVE-VOLTMETER CIRCUIT

noted that this is really a negative feed-back circuit, so that variations in the calibration due to circuit changes are minimized (though not entirely avoided).

To avoid Miller effect, which would reduce the effective

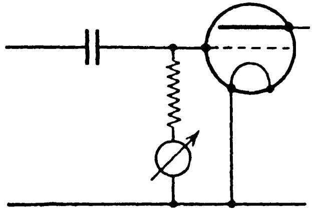


FIG. 137. CIRCUIT OF PEAK VALVE-VOLTMETER

input impedance and thus nullify one of the principal advantages of a valve voltmeter, a large condenser is shunted from anode to cathode. In Fig. 136 the small steady current through the meter is balanced out by a voltage obtained from a potentiometer across the filament, in series with a resistance *R* to avoid short-circuiting the meter.

The valve voltmeter again only measures the mean current and has to be calibrated in terms of r.m.s. voltage so that it is still subject to wave form error, but this can be minimized by reversing the leads to the input. If the reading is different the waveform is distorted, but the

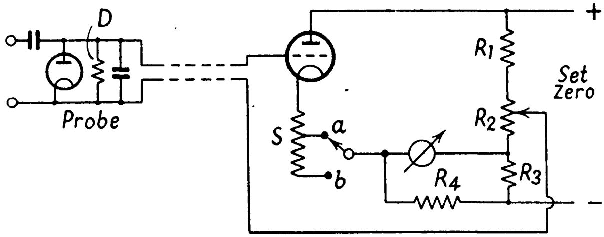



FIG. 138. DIODE VOLTMETER CIRCUIT

average of the two readings will give the correct result to a fair approximation.

Fig. 137 illustrates a special form of valve voltmeter designed to measure the peak voltage of a signal. It is, in effect, a grid rectifier with a meter in the grid circuit

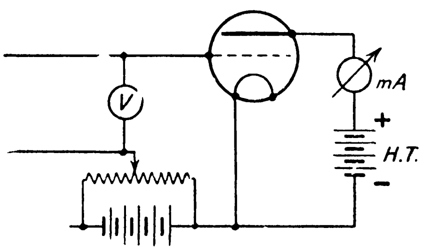



FIG. 139. CIRCUIT OF SLIDE-BACK VOLTMETER

to estimate the grid current. This current is actually proportional to the peak value of the applied signal, as is obvious from the theory of the grid rectifier. A diode could be used in place of the triode illustrated.

A development of this circuit is shown in Fig. 138. Here a conventional diode circuit is used to rectify the signal, the d.c. voltage across the diode load *D* being amplified by a cathode-follower valve. As was shown in Chapter VII,

the performance of such a valve can be made independent of the circuit values so that the calibration of the instrument can be very stable. The steady anode current through the meter is offset by a small opposing current fed from a tapping on the potentiometer $R_1R_2R_3$ across the H.T. line. The set–zero control alters the bias on the valve and so adjusts the anode current until the current through the meter is zero. Two values of cathode resistor are shown so that two different sensitivities can be obtained.

The diode itself, being small, can be accommodated with its circuit elements, in a small probe at the end of a flexible cable. This enables the actual terminals to be located quite close to the voltage to be measured, while the instrument and its power supplies, which are necessarily bulky, can be located some feet away.

Yet another type of meter is the *slide-back* voltmeter (Fig. 139). Here the bias is adjusted with no signal to give some small value of anode current. The signal is then applied, causing the anode current to increase. The grid bias is now increased until the current is the same as before. The difference in the grid bias is the peak value of the signal.

## Circuit Constants, Measurement of $Q$.

Measurement of fundamental circuit constants such as inductance and capacitance has been discussed in Vol. I. Such measurements are usually made at an audio frequency by a bridge method. A quantity which must be measured at the operating frequency, however, is the resistance, for this is invariably quite different from the d.c. value, partly because of skin effect and eddy current loss in the wire of the coil and partly because of dielectric and other losses in the remainder of the circuit.

Measurement of the h.f. resistance is both cumbersome and difficult and the more usual practice is to measure the $Q$ of the circuit. This is the circuit magnification $= \omega L/R$. The obvious method is to inject a known voltage into the circuit and measure the voltage developed across the coil when the circuit is tuned to resonance. The ratio of output to input voltage is the figure required.

There are practical difficulties in this technique. The

only convenient way of injecting the voltage is by inserting a small resistance in series as shown in Fig. 140. To avoid disturbing the circuit under test this resistance must be a small fraction of an ohm and to develop a reasonable voltage across this a large current is required. But unless the oscillator and the leads to the injection resistance are carefully shielded, voltage will be induced in to the circuit by direct induction which will invalidate the results.

Instruments known as $Q$-meters are made in which these matters are attended to. The valve voltmeter is

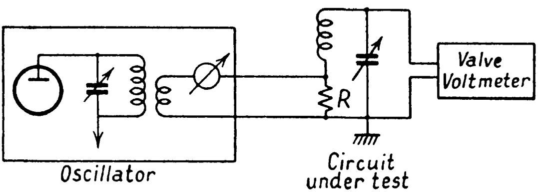



FIG. 140. CIRCUIT FOR MEASURING $Q$

built-in and calibrated directly in $Q$, and reliable measurements can be made up to frequencies of the order of 50 Mc/s.

If such an instrument is not available it is preferable to use an indirect method. The circuit under test is energized by any convenient means and the voltage developed at resonance is measured with a valve voltmeter. The circuit is then mistuned until the reading on the voltmeter falls to 0·71 times its resonant value.

The mistuning may be carried out by varying the frequency of the injected signal. Then if $f_1$ and $f_2$ are the frequencies, on either side of the resonant frequency $f_0$, at which the output is 0·71 times the maximum,

$$Q = f_0/(f_1 - f_2).$$

(This is another way of expressing the relations quoted on page 58.)

Alternatively the circuit may be mistuned by altering the capacitance. In this case, if $C_1$ and $C_2$ are the capacitances at which the voltage is 0·71 times the maximum, and $C_0$ is the resonant value,

$$Q = 2C_0/(C_1 \sim C_2).$$

## Gain Measurement.

A type of measurement often required is that of the step-up or gain of a stage. The principle adopted here is usually the same whether at high or low frequencies. The circuit is supplied with voltage at the required frequency, and a valve voltmeter is arranged to read the voltage at the output of the stage in question. The voltmeter is then

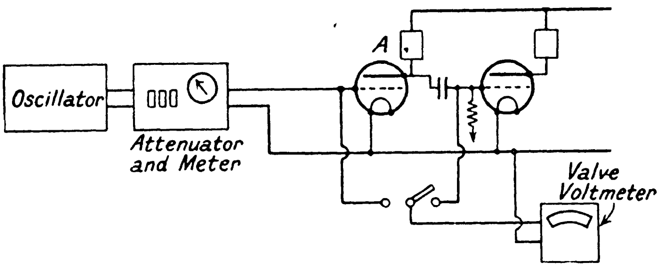



FIG. 141. TYPICAL SET-UP FOR MEASURING STAGE GAIN

transferred to the beginning of the stage, and the input is then increased till the voltmeter reading is the same as before. The increase necessary is equal to the gain of the stage.

The source of supply is an oscillator generating current of the required frequency, and the usual practice is to employ a small master oscillator with suitable provision for ensuring reasonable constancy. This is followed by an amplifying stage, the output from which passes into the attenuator. Fig. 141 shows a typical arrangement which can be used for audio frequencies measuring the gain of valve $A$. A similar layout would be used for high frequencies, with suitable modifications to allow for the higher frequency.

As in any measurement, care must be taken during the test that the introduction of the voltage at any point does not alter the operation of the circuit seriously, and that the connection of the voltage measuring device (usually a valve voltmeter) across the output or input does not in itself affect the circuit at the particular frequency at which the measurement is being carried out. If it does, some

alternative method must be devised, but the method should if possible be such that the same reading is always obtained on the indicating instrument by increasing or decreasing the input suitably. Experience shows that this is the best method, since the only variation is that of the input which can be accurately controlled.

## Receiver Measurements.

A development of this simple gain measurement is the complete measurement of the performance of a receiver. Here a small dummy transmitter is employed known as a *signal generator*. It usually consists of an oscillator followed by a modulating stage (the modulation usually being employed on an amplifying valve instead of on the oscillator itself to avoid frequency modulation as explained in Chapter I) and a radio-frequency attenuator. This provides a source of modulated voltage which can be varied in output from about a tenth of a volt down to a few microvolts. It is obvious that the most careful shielding must be adopted to ensure that energy can only be obtained from the generator **via** the output terminals, where it is under proper control, and not by any direct radiation which would, of course, invalidate the measurement. In any case, it is always desirable to keep the receiver under test several feet away from the generator, to be on the safe side.

Receiver measurements are carried out by impressing on the aerial a signal having a standard modulation (usually 30 per cent at 400 cycles), and adjusting the input so that when the receiver is properly tuned and operating at full efficiency the output is 50 mW. The power output is measured by a suitable (thermal) mêter in series with a resistance equal to the optimum value for the particular output valve in use, a choke or transformer output stage being employed to keep the d.c. out of the meter.

This measurement is carried out at various frequencies within the range of the receiver and the sensitivity curve obtained. Fig. 142 shows a typical layout.

Selectivity is measured by mistuning the oscillator (leaving the receiver unaltered) and increasing the input

until the receiver again gives 50 mW. output. This can be done for a series of frequencies gradually getting farther away from the tune of the receiver, and a selectivity curve plotted. It will, in fact, be an inverted resonance curve.

Fidelity measurements are taken by tuning the receiver to some suitable radio frequency and then varying the frequency of the modulation, keeping the percentage constant. The output power is then noted. If the receiver were perfectly faithful, of course, the output power would

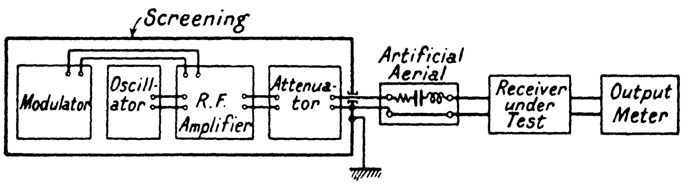



FIG. 142. SET-UP FOR MEASURING RECEIVER PERFORMANCE

be the same irrespective of the modulation frequency. In practice, this is very far from being the case.

These are the standard measurements made on a telephony receiver. For communication work, of course, the fidelity curve is not usually required. Numerous other measurements can be made. 50 mW., for example, does not represent full power output by any means, and one can examine the behaviour of the receiver as the input is gradually increased, or the relative output with constant input but differing depths of modulation. Special tests are also made using two signal generators to determine the degree of cross-modulation present (see page 71). One generator is adjusted to give normal output and its modulation then switched off. The other generator is then mistuned by a certain amount and modulated in the normal manner. By cross modulation, this modulates the wanted carrier and the input of the interfering signal is adjusted until normal output from the wanted carrier is obtained. This is then repeated for varying amounts off tune, the interfering signal input becoming progressively larger as it becomes more and more mistuned.

## Modulation Measurement.

An important factor in these measurements is the estimation of the modulation, and various methods are employed here. Where it is not possible to obtain access to the low-frequency modulating voltage, the modulation depth may be estimated by noting the change in the operating current when the circuit is modulated. Although the mean value of a modulated current is the same as the normal carrier wave, the r.m.s. value increases, the rise being approximately 23 per cent with full (sine wave) modulation. Actually, the depth of modulation is given by the expression

$$M = \sqrt{\left[2\left(\frac{I_2^2 - I_1^2}{I_1^2}\right)\right]} \times 100 \text{ per cent}$$

where $I_1$ is the unmodulated current, and $I_2$ is the modulated current.

In a signal generator, however, it is usually practicable to measure the actual audio-frequency voltage across the modulating choke, and the peak a.f. voltage divided by the h.t. voltage gives the percentage modulation rather more accurately than can be deduced from the current rise.

Still another method is to apply the modulated signal to a detector valve and measure the change in d.c. anode current and the alternating anode current. The latter may be estimated by measuring the modulation frequency voltage across a resistance in the anode circuit. The a.c. component divided by the change in d.c. current gives the modulation depth.

The diode circuit of Fig. 44 can also be used by inserting a microammeter in series with $R$, and also measuring the l.f. voltage across $R$ with a valve voltmeter (a h.f. choke being inserted in the voltmeter lead to keep out the h.f.).

Then $M = \dfrac{(\sqrt{2})\beta V_{lf}}{I_R R} \times 100$ per cent, where $\beta$ is the rectification efficiency of the diode.

A visual check on modulation depth can be made with an oscillograph such as is described in the following section. This enables the modulated wave-form to be displayed and the height of the peaks and the troughs can be measured.

If these are heights $a$ and $b$ respectively, the modulation depth is $(a - b)/(a + b)$. It is also possible, by a modified form of connection, to examine not only the extent but also the linearity of the modulation. This technique is described in *Cathode Ray Oscillographs* (Pitman) by the present author.

## The Cathode-ray Tube.

Reference should be made in conclusion to the cathode-ray tube which is taking its place as a recognized measuring

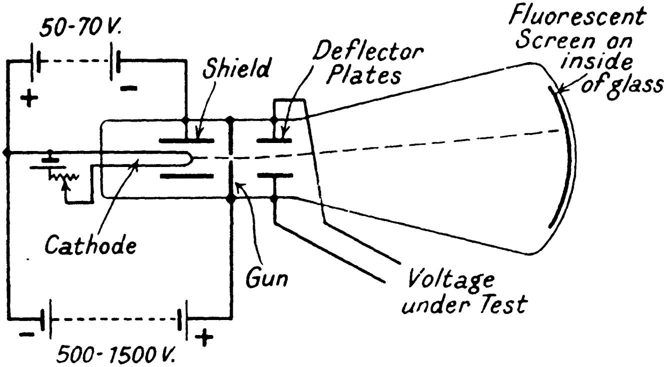



FIG. 143. BASIC CIRCUIT OF CATHODE-RAY OSCILLOGRAPH

instrument in communication work. This consists, in essence, of a cathode emitting electrons, and an anode in the form of a plate with a hole in the centre which is maintained at a steady positive potential. Some of the electrons emitted from the cathode shoot through the hole in the anode in the form of a pencil or stream of electrons which is allowed to impinge on a fluorescent screen at the end of the bulb. The point where the electron beam strikes the screen is evidenced by the appearance of a small spot of light, usually of a greenish or bluish colour depending upon the material of the screen.

In order to concentrate as many of the electrons in the beam as possible a *cylinder* or *shield* is placed round the cathode, this being connected to a negative potential rather

like the grid of an ordinary triode. This prevents the electrons emitted by the cathode from dispersing, so that the greater concentration of electrons in the centre results in the majority of them finding their way through the hole in the anode, or *gun*, as it is often called. The ordinary oscillograph tube uses 500 to 1 500 volts on the anode and 50 to 70 volts on the shield.

The electron beam may be deflected by electrostatic or magnetic means. By placing a pair of plates on either side of the beam and applying to these plates suitable voltages, the spot of light may be deflected from side to side. A similar pair of plates set at right angles will cause a deflection in a vertical direction, and between the two a composite motion can be obtained which is very useful for analysing wave forms and similar phenomena.

### Wave-form Analysis.

For example, the application of an alternating voltage to the vertical deflecting plate will cause the spot of light to lengthen out into a thin vertical line. If, at the same time, a voltage of lower frequency is applied to the horizontal plates, the vertical movement of the spot will be spread out so that a wave form will be traced out on the screen.

The horizontal deflecting voltage should obviously be of rather special character if the best use is to be made of this phenomenon. If the voltage slowly and regularly increases until the spot has been deflected over the full width of the screen, and then rapidly returns to zero and commences the process again, it is possible to cause the successive traces of the wave to lie one on top of one another so that a stationary picture of the actual wave form is obtained. The only conditions required are that the time taken for the spot to travel across the screen shall be an exact sub-multiple of the frequency, and this is easily arranged.

### Time-bases.

The *time-base*, as it is called, which gives the deflecting motion is usually some form of condenser-charging circuit.

For instance, the charging of the condenser in a self-oscillating circuit and the subsequent discharge could be used for the purpose, but it is more usual to employ what are known as *gas discharge* tubes for the purpose.

These are the triode valves containing a small quantity of gas. If the grid of the valve is maintained at a negative potential, no current will flow in the anode circuit until the anode voltage exceeds a certain amount. The arrangement is like an anode bend detector biased to cut-off point.

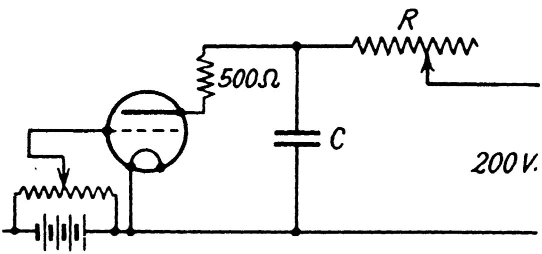



FIG. 144. TIME-BASE CIRCUIT FOR CATHODE-RAY TUBE

When the anode voltage exceeds the critical value, however, current will flow, and this current will immediately ionize some of the molecules of gas, liberating further free electrons (and also heavy positive ions which travel towards the cathode). The freed electrons, in turn, produce further ionization and a very rapid increase of current is obtained—sufficient to damage the valve unless a limiting resistance is inserted in the anode circuit.

The way the tube is used is shown in Fig. 144. The condenser is charged through a resistance, and acquires its charge relatively slowly, depending upon the values of condenser and resistance. The larger we make either of these two, the slower is the charge. As soon as the voltage on the condenser reaches a critical value, the gas discharge tube becomes conducting and the condenser discharges very rapidly indeed, after which the process re-commences.

The frequency of the operation depends upon the values of resistance and capacitance and the grid bias of the gas

discharge tube relative to the applied voltage. This form of discharge will be seen to give what we require, namely a steady increase in voltage coupled with a very rapid discharge and re-commencement. The charging of a condenser through a plain resistance, however, is not uniform, which gives rise to some distortion if the wave form is being examined critically. To avoid this a pentode valve may be used instead of a resistance as shown in Fig. 145. The

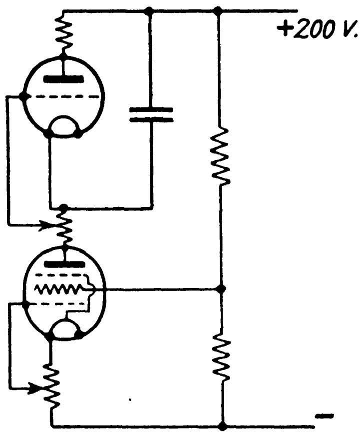



FIG. 145. CONSTANT-CURRENT TIME-BASE CIRCUIT

anode current of a pentode is practically independent of the voltage applied, depending mainly upon the voltage on the control grid of the valve. The device is, therefore, capable of being used as a variable resistance, and it will charge the condenser at a constant rate.

Other forms of time base have been evolved, but it is unnecessary to detail them here as they operate on similar principles.

**Focusing.**

It is necessary to focus the beam of electrons on to the screen so that they give a sharp image. It is usually done in small tubes by introducing a certain amount of gas into the bulb. The electrons in the beam encounter occasional

molecules of gas, ionize them, and release a few free ions. These are attracted to the beam, which is, of course, negatively charged, and they in their turn act as a concentrating force on the various electrons in the beam, which all tend to bunch together in a pencil. The actual focusing effect

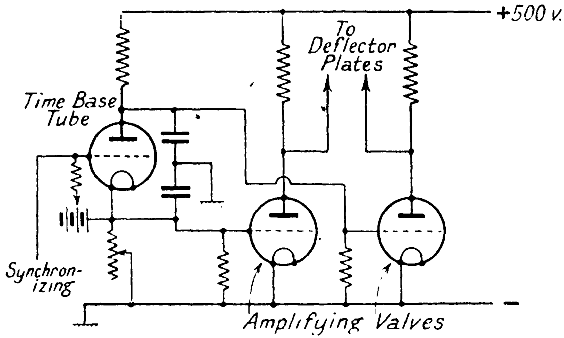



FIG. 146. AMPLIFIED PUSH-PULL TIME BASE

depends upon the current in the electron beam, which is controlled by altering the shield voltage.

## Hard Tubes.

An alternative method of focusing, which does not involve the introduction of gas, is to use two anodes, one beyond the other, and to adjust the voltages on the shield and two anodes so that a concentration is obtained. The focusing action here is dependent upon the relative sizes and positions of electrodes and is, of course, due to the electric field produced thereby. Tubes of this type are called "hard" tubes in distinction to the "soft" gas-filled tubes. They require somewhat higher anode voltages of the order of 2 000 to 5 000 volts.

The advantages of the hard tube, which has now superseded the soft variety, are longer life due to the avoidance of bombardment of the cathode by heavy gas ions, and relative independence of the focus and beam current. This enables the shield to be used for controlling the intensity of the spot, which is very valuable in television.

The tubes, however, are somewhat less sensitive and are liable to defocus if the voltage on the deflector plates is large. This defect may be minimized by using a push-pull time base comprising two valves driven off the normal time-base circuit as shown in Fig. 146. This arrangement has the advantage of requiring only a small time base, which is easier to make.

## Magnetic Deflection.

It is possible to deflect the beam by applying a magnetic field at right angles to the beam. The direction of deflection

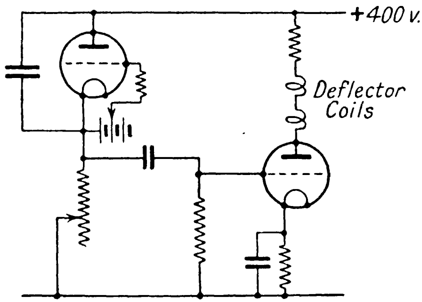



FIG. 147. CIRCUIT FOR PROVIDING MAGNETIC DEFLECTION

of the beam is then at right angles to the direction of the magnetic field, and by arranging two pairs of coils at right angles to one another we can produce the desired deflection exactly as in the electrostatic case. The disadvantage of this method of working is that power is consumed by the coils.

The coils must be fed from an amplifier valve as shown in Fig. 147, this valve being driven by the time base and arranged to operate over a linear portion of the characteristic so that the current is directly proportional to the grid voltage.

The possibility of deflecting the beam magnetically must always be borne in mind, because stray magnetic fields from nearby apparatus may often distort the wave form and give very misleading results. The cathode-ray tube

should be at least 3 and preferably 6 ft. clear of any electrical apparatus, particularly mains transformers, unless such apparatus is heavily shielded with material of high permeability such as Mumetal. Alternatively the neck of the tube itself may be enclosed in a Mumetal sheath.

This is a very brief review of an extremely important subject. For further detailed information the reader should refer to *Cathode Ray Oscillographs* (Pitman), by the present author.

# CHAPTER XV

## PICTURE TRANSMISSION AND TELEVISION

THE transmission of pictures by wire or radio link is in everyday use, while television has emerged from the experimental stage and can now take its place with standard engineering developments in radio communication. The difficulties in its exploitation are commercial and economic rather than technical.

The basic method of transmitting either a still picture, for press or similar purposes, or a succession of pictures at very short intervals, to constitute a television image, is carried out by a process known as "scanning."

### Scanning.

The picture is divided up into a series of thin strips or lines, and the point under examination moves along each line in turn. Each individual line is thus composed of a number of successive elements which may be light, dark, or half-tone, according to the nature of the image, and if this light and shade can be translated through a photo-electric cell into electric currents it is possible to obtain a varying voltage having a modulation which follows the variations of light and shade in the original strip of the picture.

Having analysed one line, the transmitter then proceeds to an adjacent line and so on until the whole picture has been covered. In this way the picture or image is divided up into a series of successive picture elements, each one of which will produce its corresponding modulation voltage, and clearly such currents can be transmitted either by wire or by a radio link to a distant point.

At the receiving end it is necessary to re-convert the current into light, the intensity of which is directly proportional to the current at each particular instant. The light must be focused on to or otherwise caused to illuminate the receiving screen in a similar series of lines and in

synchronism with the transmitter. Thus the spot of light at the receiving end will travel along one line and the intensity of the illumination at each successive point will be controlled by the current received from the distant end, so that a series of light and dark picture elements are built up. The spot then proceeds with the second line and so on until the various picture elements have all been re-assembled and a complete picture is obtained which should be a duplicate of the original scene.

With picture telegraphy the whole process can take several minutes since only one copy is required. With television, on the other hand, the whole picture must be built up in a small fraction of a second and must then be repeated over and over again, the picture frequency being sufficiently rapid to enable the visual persistence of the eye to blend the whole impression into a continuous image as in the case of the cinema.

## Photocells.

The process at the transmitter depends essentially on the conversion of light into electric current. This involves the use of some form of photo-electric cell (or photocell). Certain substances, such as cæsium*, emit electrons under the influence of light in a similar manner to the emission of electrons from a hot cathode. A photocell therefore consists of an evacuated glass bulb containing a cathode covered with cæsium with an anode close by, maintained at a positive potential. The external circuit contains a high resistance and the passage of the electrons through this develops a voltage which is proportional to the light falling on the cathode.

Since the internal resistance of the cell is nearly infinite, the larger we can make the load resistance the greater the sensitivity. Above a small threshold value the actual anode voltage has no appreciable effect on the emission and as the emission is measured in microamperes it is customary to use load resistances of the order of one megohm. The limiting factor is the frequency of the light

* Pronounced seesium.

variations, for as this increases, stray circuit capacitances begin to reduce the effective load impedance and the value of anode resistance has to be restricted.

In a practical photocell the cathode is often deposited direct on the inside of the glass, which is usually silvered first and then coated with the photo-sensitive material since this increases the sensitivity The anode is usually a small ring to serve as an efficient collector without obscuring the light. Rubidium and potassium are also used for the cathode. They are less sensitive than cæsium but respond better to yellow and red light.

For low frequency applications the sensitivity can be still further improved by introducing a small amount of inert gas into the cell. The primary electrons emitted by the cathode encounter molecules of gas which they ionize by collision and the total emission is increased many times. With such a cell, however, the anode voltage also affects the emission and it must not exceed a certain critical value or a continuous discharge occurs which damages the cell. Moreover after each light impulse the gas must have time to de-ionize so that the arrangement is unsuitable for frequencies above a few thousand c/s.

Photo-electric emission, however, can be increased by the use of electron multipliers as explained on page 259.

## Facsimile Transmission.

For the transmission of pictures by wire or radio, the original is scanned by a source of light operating through some convenient mechanical arrangement, which may be a mirror drum system similar to that described for television later, or the simple arrangement of wrapping the picture round a drum, causing the drum to rotate and at the same time tracking the light axially down the picture. This will obviously scan the picture in the manner described and a somewhat similar arrangement can be used at the receiving end.

For the reception photo-sensitive paper is usually employed so that the variation in the intensity of the light spot affects the paper in the same way as the varying light transmitted through a photographic negative, and when

the whole picture is developed the image comes up in the usual way and is then fixed.

There are numerous matters of technique which can only be referred to briefly here. In particular, since the whole transmission takes some minutes, the frequencies of the signals involved are very low. This would necessitate d.c. amplifiers at the transmitting end to augment the very small voltages produced by the photocell, and it would also demand characteristics from the line over which the transmission was sent far in excess of those which are satisfactory for normal telegraphy or telephony. It is customary, therefore, to chop up the impulses by an interrupter, usually interposed in the path of the light beam itself, so obtaining a rapid succession of impulses, the amplitude of which is varying more slowly in accordance with the light and shade of the picture points of the image being scanned.

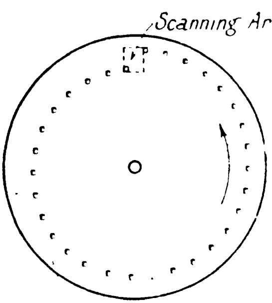



FIG. 148. SCANNING DISC FOR TELEVISION

It is, of course, necessary to ensure that the mechanism at the receiving end revolves at exactly the same speed as at the transmitting end. This is ensured by the transmission of synchronizing signals in somewhat similar manner to that adopted in television practice.

## Television Scanning.

The method of scanning the scene or image in a television transmitter is slightly different. It may be effected by mechanical or electrical means, the simplest of the former being the *Nipkow* disc (Fig. 148), named after the Polish scientist who first suggested it. This is a disc having a series of holes arranged on a spiral. As the disc rotates, the holes traverse the scanning area in a vertical or horizontal direction so that the image is explored in a succession of lines in the required manner. The subject can be illuminated and

an image focused down with a lens on to the face of the disc, while behind the disc is situated a photocell which converts the light impulses into electric current.

This usually requires a very large illumination owing to the extremely short time during which the photocell is influenced by any particular picture point, and an alternative method developed by the Baird Company is to locate a powerful source of light behind the disc and to project this on to the image to be scanned. In this way a spot of light travels over the subject in the same series of lines and the light reflected at any instant is picked up and converted into electrical current by banks of photocells suitably located. This arrangement, known as the *flying-spot system*, is much more pleasant to the performer.

## Mirror Drum.

An alternative method of mechanical scanning is by the use of a mirror drum. A cylindrical drum is provided with a series of mirrors around its periphery and a source of light is projected on to the mirrors, reflected off them and focused through a lens on to the subject to be scanned. The rotation of the drum causes the spot to move radially as required, while the successive mirrors are slightly staggered relative to one another so that the successive spots of light trace out linear paths adjacent to one another. Various modifications of this basic idea have been used, notably by Mihaly, who employed a ring of stationary mirrors with a rotating mirror in the centre, thus achieving the same result with a very light moving mechanism.

## Mechanical Receivers.

Reception can be carried out at the distant point by the use of mechanical receivers of similar construction. In place of the photocell a source of light has to be used. For example, a neon lamp placed behind a disc can be caused to glow more or less brightly in accordance with the received signal, so that at any instant when the transmitter is traversing a bright picture point the lamp at the receiver will glow brightly and will produce a bright picture point in the receiving mosaic. Hence, once again

the image is built up, the only requirement being that the apparatus at transmitting and receiving ends shall run in synchronism.

A mirror drum receiver can be used equally well and can indeed be employed to project the picture on to a reasonable-sized screen. The difficulty here, however, is to arrange a suitable source of light, for a neon lamp only gives a feeble illumination, and any form of filament lamp is unsatisfactory because the thermal inertia of the filament prevents the light from being modulated sufficiently rapidly.

## Light Modulation.

To obtain any satisfaction, therefore, it is necessary to use a powerful light source and devise some external method

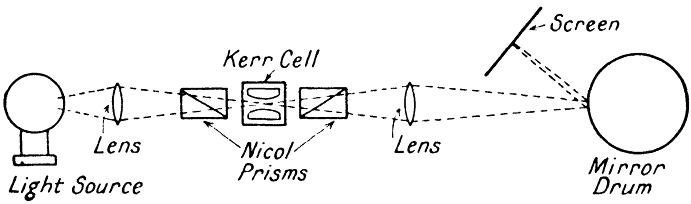



FIG. 149. PRINCIPLE OF MIRROR DRUM RECEPTION WITH KERR CELL LIGHT MODULATION

of modulating its intensity. One method, suitable for low and medium definitions, uses the polarization of light. Ordinary light contains vibrations in all directions at right angles to the direction of travel. With polarized light the vibrations are only in one particular plane. To the normal eye polarized light appears exactly the same as ordinary light, but it has certain peculiar properties which show up under abnormal conditions.

Polarized light is usually produced by a device known as a *Nicol* prism, which is an arrangement of two prisms of Iceland Spar or calcite. The particular plane of polarization of the light depends upon the position of the Nicol prism, and as it is rotated so the plane of polarization of the emergent light rotates.

Clearly, if we have two such prisms in succession so adjusted that their planes of polarization are at right angles, no light will pass through the system. The first

prism, known as the "polarizer," will polarize the light in one plane, but if the second prism, called the "analyser," is set at right angles to this direction it will not accept any light polarized in this plane and the result will be darkness. If the second prism is rotated, a certain amount of light passes through, reaching a maximum when its plane of polarization is the same as that of the first polarizing prism.

### The Kerr Cell.

It is impracticable to rotate the prisms mechanically at the very high frequencies required, but it is possible to achieve the effect electrically by using a Kerr cell, which is a series of plates arranged on either side of the light beam and connected to a source of voltage. The electrostatic field has the effect of twisting the plane of polarization.*

The polarizer and analyser are set at right angles so that no light passes through. The application of voltage across the Kerr cell electrodes then causes the effective plane of polarization to rotate and allows some light to pass through the prism exactly as if the prism itself had been rotated. This operation can be carried out electrically many hundreds of thousands of times per second, and therefore provides a very satisfactory method of modulating the light, which can therefore be made extremely powerful, giving a good picture.

### Supersonic Light Valve.

For high definitions, the self-capacitance of the Kerr cell precludes satisfactory operation, and use is made of a different method of modulation depending upon the

* Strictly speaking, the electric field in a Kerr cell splits up the polarized ray into two component rays at right angles to one another, having differing intensities depending upon the strength of the field. As the field is varied, the intensities of the two component rays relative to one another change, producing what is known as "elliptical polarization." The exact mechanism is described in greater detail in books on Television or Optics, notably *Television, Theory and Practice* by the author, and the fact is mentioned here for the sake of accuracy. In a simple exposition of the subject, however, it is sufficient to consider the Kerr cell as producing a rotation of the plane of polarization.

interference effects produced by transmitting supersonic waves through a column of liquid. The action is complex, but is being used successfully for the production of large screen images. For further details, refer to *Television, Theory and Practice* (Chapman and Hall), by the present author.

**Cathode Ray Television.**

The development of the cathode ray tube gave television a considerable impetus. Mechanical methods become increasingly complicated as the number of lines is raised. The British Television Commission adopted a standard of no less than 405 lines, while American workers are using still higher definitions. Obviously, the detail in a picture increases rapidly with the number of lines for this means that the picture elements themselves are proportionately smaller, and the detail improves in the same way as a fine screen half-tone reproduction on good paper is considerably better than a coarse-screen newspaper block.

The advantage of the cathode ray tube is that the electron beam (the production of which was described in the last chapter) can be deflected electrically at exceedingly high frequencies owing to its entire absence of inertia. The method of deflection is the normal one employed for oscillograph practice, i.e. the use of a suitable time-base such as that illustrated in Fig. 144. This causes the spot of light to move across the screen, fly back rapidly, and recommence its travel. If a similar deflecting voltage of much lower frequency is applied to the second pair of deflector plates, the spot will start its second line slightly displaced from its original position, so that two lines will run parallel and so on for the third and successive lines until a complete framework has been obtained. The voltage on the low-frequency plates can then collapse and the process will start again building up a second framework of lines on top of the first, and repeating this indefinitely at the rate of 50 per sec. or whatever the picture frequency happens to be.

The number of lines in the framework is clearly dependent upon the ratio of the frequencies of the two time-bases. If the high-frequency time-base produces 6 000 lines

per sec., and the low-frequency time-base operates 25 times per sec., there will be 6 000/25 = 240 lines in the picture. The usual modern method of scanning is for these lines to be horizontal with the picture deflection vertical, although the earlier 30-line television used vertical lines with an horizontal picture deflection.

The variation of the intensity of the light is controlled by altering the voltage on the shield. The ordinary soft tube is not suitable for this purpose, for variation of shield voltage in such a tube merely controls the focusing and, although it will reduce the intensity of the spot to some extent, this is usually accompanied by serious de-focusing. Special hard tubes have, therefore, been developed, as explained in Chapter XIV, in which the variation of shield volts merely alters the intensity without appreciably upsetting the focus, this latter factor being controlled by adjustment of the voltage on the various anodes.

## The Iconoscope.

In view of this flexibility, it is not surprising that attempts have been made to utilize the tube for transmitting apparatus. After years of experiment, some satisfactory arrangements have been evolved. One is the *Iconoscope* of Zworykin. In this arrangement the usual fluorescent screen at the end of the oscillograph tube is replaced by a plate containing a mosaic of very small photocells. The ordinary photocell contains an emitting surface of cæsium on a thin film of silver which is deposited on the glass or emitting surface first. By suitable manufacture it is possible to deposit this emitting surface in a series of very small globules, each one of which therefore becomes an individual and independent photocell. This mosaic is deposited on but insulated from a suitable metal plate and is situated at such an angle that it can be viewed conveniently through the side of the tube, as shown in Fig. 150.

The scene to be televized is focused down on to this plate, and each of the tiny photocells in the mosaic will thus respond individually to the amount of light in that particular part of the scene. Since there is, as yet, no circuit, the cells will acquire a charge proportional to the intensity of the light.

A cathode-ray beam generated in the usual way is now caused to scan over the plates by deflecting voltages of the usual type and as it reaches each individual photocell in the mosaic the latter is discharged, producing a current in the external circuit proportional to the charge.

This method, therefore, combines the flexibility of the cathode-ray tube with the necessary photo-electric properties

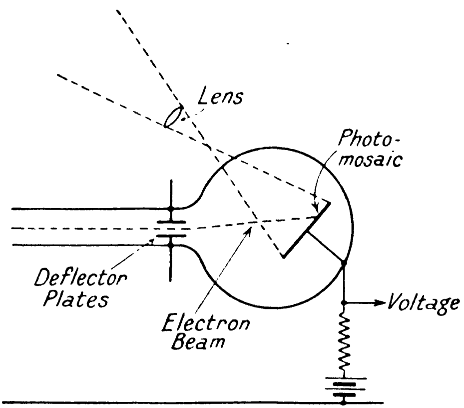



FIG. 150. PRINCIPLE OF ZWORYKIN ICONOSCOPE

required to convert light to electric current and it has been found practicable to scan with a definition approaching 500 lines.

**The Electron Camera.**

An alternative method has been developed by Farnsworth and is known as the electron camera. In this arrangement the photo-electric surface is in the form of a flat plate on which the image to be televized is focused. In front of the plate is an arrangement of anodes of cylindrical form, so arranged that they attract electrons from the photo-electric cathode in straight lines. The emission from each part of the cathode is proportional to the light falling on that particular portion, so that there is emitted from the cathode a stream of electrons the intensity of which at any point

in the cross-section is dependent on the light in the image focused on the cathode.

At the far end of the tube is a small collector electrode so enclosed that only a small point faces this composite electron beam, and the beam as a whole is then deflected in vertical and horizontal directions by scanning voltages,

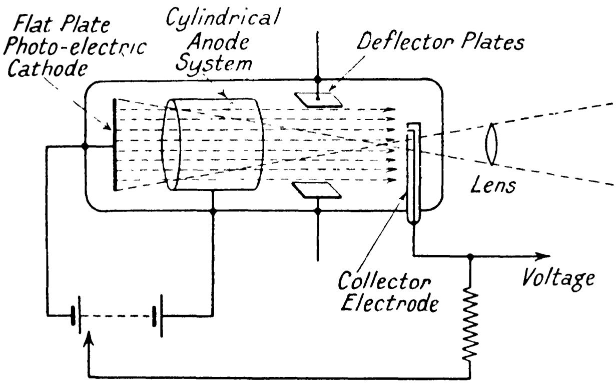



FIG. 151. ILLUSTRATING ACTION OF FARNSWORTH ELECTRON CAMERA

just as with the much simpler beam of an ordinary cathode ray tube. Hence, each part of the composite electron stream passes the collector electrode in turn, and as it does so it passes a current through the external circuit proportional to the intensity of the beam at that point.

Once again, therefore, it will be seen that the light value in the original image has been converted into electric current by an arrangement having all the flexibility of the cathode ray tube, and this method has been used to produce pictures containing as many as 700 lines.

## Electron Multipliers.

The great difficulty in television is to obtain sufficient light on the photocell, for the amplification which can be used is limited by the background noise which is produced due to electronic disturbances as mentioned on page 68. Considerable improvement has been possible as the result

of the development of electron multipliers which are devices using secondary emission.

A simple form of multiplier is shown in Fig. 152. The electrons emitted from the photo-electric cathode are focused on to a plate which is designed to be particularly susceptible to secondary emission. This means that if it is struck by electrons moving with sufficient rapidity it will in turn

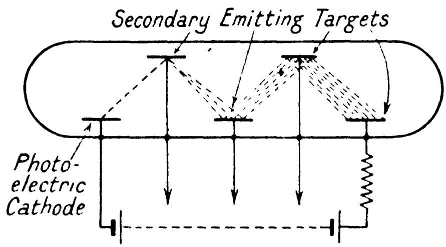



FIG. 152. SIMPLE ELECTRON MULTIPLIER

emit electrons of its own in the proportion of perhaps five or ten electrons for each one which strikes it. Hence, there will be emitted from this secondary plate a current five or ten times as great as the original current, and this process can be repeated by successive emission from subsequent plates all arranged at a higher potential than the preceding one until an amplification of hundreds of thousands of times is obtained in the one tube.

Since the circuit contains no coupling impedances or subsidiary source of emission, the question of background noise does not appear and very weak illuminations on the photocell cathode are capable of giving some milliamps emission at the far end.

It is indeed conceivable that this development will radically alter the trend of ordinary radio-receiver design, and secondary emission valves having about twice the usual slope have already appeared on the market.

By the use of electron cameras or iconoscopes coupled with electron multipliers of this form, it is possible to have television cameras which are as sensitive as an ordinary still camera, and which can be merely pointed in the direction of the particular scene, using ordinary daylight

(or artificial light) for the illumination and to obtain multi-line pictures without difficulty.

**Synchronism.**

The synchronism of the received picture is not by any means the difficult problem it might appear. At the end of each line a special synchronizing impulse is sent out and a similar picture-frequency synchronizing impulse at the end of each frame. Moreover, these impulses are stronger than the strongest modulation. The usual procedure is to modulate with the picture frequencies from 100 per cent down to 30 or 40 per cent of the peak carrier, while the synchronizing impulse momentarily causes the carrier to drop practically to zero.

It is possible, therefore, to use over-biased valves in the receiver which are unable to respond to any of the normal modulation signals but will pass current on the arrival of the extra strong synchronizing pulse. In this way the synchronizing signals are completely separated from the modulation.

In cathode ray reception the synchronizing impulses are applied to the appropriate time bases, causing them to trigger at the correct instant. It was explained in the last chapter that the discharge of the time base is controlled by the grid voltage on the discharge valve. Clearly, if the circuit is in a condition approaching the critical value, the application of a sharp pulse to the grid of the valve will trigger it. Thus, the actual discharge point is controlled by the received signal itself.

With mechanical systems it is usually only necessary to use the line-frequency synchronizing signal and cause this to control the speed of the motor by some suitable means. One way of doing this is to have a phonic wheel on the motor shaft containing a number of teeth which pass close to an electromagnet. This magnet in turn is energized by the synchronizing signal. If the tooth is exactly opposite the pole piece at the instant of arrival of the synchronizing impulse (as it should be), no effect is produced, but if it is slightly behind or ahead, the synchronizing impulse will produce a magnetic field tending to pull the motor into step.

## Frequencies Involved.

The frequencies involved in television technique are exceedingly high. The finest pattern which can be reproduced can be considered as being made up of alternate black and white squares each having the dimensions of the scanning spot. The width of the scanning spot will be equal to the width of the line, so that if there are $n$ lines it is easy to show that the frequency of such a mosaic would be

$$fpn^2/2 \text{ c/s.}$$

where $p$ is the ratio of picture length in the line direction to the depth across the lines, and $f$ is the picture frequency in pictures per sec.

With a 405-line 5/4 picture repeated 25 times per sec. this works out at 1·52 mc/s.

Hence, no ordinary broadcast technique will do. It is necessary in the first place to use a very high frequency carrier wave, and for this reason wavelengths of the order of 7 metres are being employed. As explained in Chapter XIII such wavelengths only have a limited range except under abnormal conditions, and the proposal therefore is to set up a chain of stations all over the country, each serving its own district.

On a wavelength of 7 metres, corresponding to 42·8 mc/s, a modulation of one or even two mc/s is still only a small fraction of the carrier wave and therefore practicable.

## Television Receivers.

The television receiver, of course, requires particular attention. The aerial system has to be of the special type discussed in Chapter XIII, while in order to cover the very wide modulation frequency, special methods have to be used. Direct amplification at these very high frequencies is extremely difficult, although it is done commercially.

An alternative method is to use the superheterodyne principle and to convert the incoming signal to a frequency of about 20 mc/s. It is possible to obtain reasonable amplification at such frequencies comparatively easily. In order to obtain the very wide band spread required, band-pass filters are used in the intermediate stages, these being

somewhat overcoupled and sometimes even shunted with resistance in order to give a very wide spread. Both these methods involve a loss of amplification and a gain of between 5 and 10 per stage is all that can be obtained. However, by using a chain of such valves in cascade satisfactory amplification is achieved.

If any amplification is required subsequent to the second detector, this again must be very carefully designed. Resistance coupling is the only satisfactory method and the constants of the circuit must be so chosen that they receive well down to the picture frequency (usually 25 per sec.) and well up to the modulation frequency of 1 mc/s or more. The latter requirement involves the use of screen grid or pentode valves owing to the reduced Miller effect obtainable, but even so the residual capacitance of 20–30 $\mu\mu$F., necessitates the use of very low anode resistances of the order of 1 000–5 000 ohms, so that a gain of 5 to 10 per stage is all that can be conveniently obtained.

## Effect of Inadequate Response.

Good low frequency response is essential for adequate reception of sustained black or white areas. Theoretically, a sustained tone of any intensity requires the transmission of an unvarying signal, which can only be achieved with a d.c. amplifier. A satisfactory approximation, however, can be obtained with a resistance-coupled amplifier of sufficiently long time constant and if the design is so arranged that it gives a negligible phase shift—not more than about 10 degrees over the whole amplifier—at the lowest frequency, satisfactory reception will result. The picture is always broken up somewhere so that the lowest frequency can be taken as the picture frequency, usually 25 or 50 cycles per sec.

High frequency response is obviously necessary for the transmission of adequate detail and if the upper frequencies are lacking the picture appears out of focus.

## CHAPTER XVI

# POWER SUPPLY CIRCUITS

Both receivers and transmitters are now very largely operated from the a.c. supply mains. This involves the use of circuits to convert the a.c. to steady d.c. at the required potential. The transformation is effected by means of one or more rectifiers, together with suitable networks for storing the d.c. and smoothing out any irregularities arising from the pulsating nature of the output from the rectifiers.

Fig. 153 illustrates four simple types of circuit. In the first circuit the mains supply voltage is transformed up or down as required and the voltage on the secondary is applied through a rectifier to a condenser $C$. This condenser receives a pulse of current every time the point $A$ becomes positive so that a charge is built up which is then dissipated through the load across the circuit. Conditions adjust themselves until the current drawn from the transformer in the form of pulses is just sufficient to replenish the loss of charge during the idle portion of the cycle.

Fig. 153 (*b*) shows a full-wave circuit in which the secondary of the transformer is centre-tapped and two rectifiers are used, connected one from each extremity. During the half cycle which makes the point $A$ positive the top rectifier conducts, while during the succeeding half cycle the bottom rectifier conducts. The current is fed into a single condenser $C$ so that the charge is replenished twice as quickly, which results in a higher output and less variation in the voltage on the condenser.

Fig. 153 (*c*) shows a bridge rectifier, an arrangement which avoids the use of a centre-tapped transformer by employing four rectifiers in a bridge formation. The current flows first through $PQ$, through the condenser $C$ and back through $SR$. On the succeeding half-wave it flows via $RQ$, $C$ and $SP$. The performance of this class of circuit is similar to

that of 153 (*b*). It has the advantage that only a single secondary winding is required and, as will be seen later, the transformer losses are reduced. There is, however, a small additional rectifier loss since there are two rectifiers in series on each channel.

Fig. 153 (*d*) shows a voltage doubler circuit. This is an arrangement employed where it is desirable to restrict the voltage on the transformer secondary, or where, for some

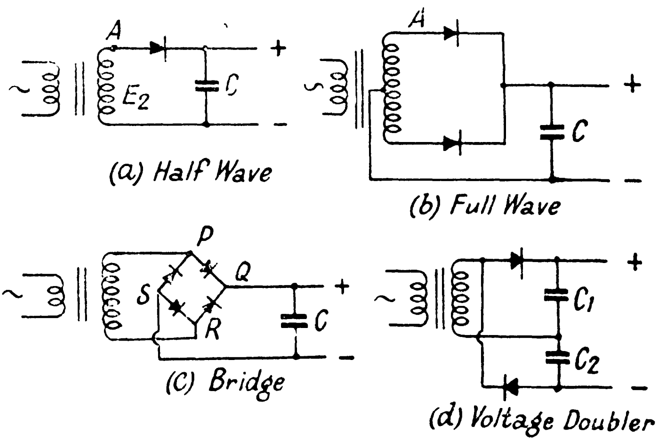



FIG. 153. FOUR MAIN TYPES OF RECTIFIER CIRCUIT

reason, there is only a limited voltage available. It gives approximately twice the voltage at the output as is obtained from Fig. 153 (*c*). Current flows through the top rectifier, and back through the condenser $C_1$. The next half cycle flows through the condenser $C_2$ and back through the bottom rectifier. Each of these pulses charges the appropriate condenser, so that the total voltage applied across the load is the sum of the two, which is thus twice the normal voltage. The two condensers $C_1$ and $C_2$ are in series so that the effective capacitance is halved, in consequence of which, as will be seen later, the regulation is liable to be poor. This class of circuit, therefore, is mainly used for the supply of high voltage at a low current.

## Ripple Voltage.

Let us, at the outset, confine our attention to the first two circuits. It should be made clear that the rectifiers used may be of any type, either valve, metal or chemical. (Symbolic rectifiers have been illustrated to avoid any complication of the circuit by the introduction of filament windings on the transformers, etc.) Considering circuit 153 (*a*) it will be clear that, after the first few cycles, there will be a voltage on the condenser $C$, of which the instantaneous value is fluctuating slightly, falling as current is taken from the condenser by the load during the idle period of the cycle and rising sharply during the period over which the charge is fed into the condenser again through the rectifier. This fluctuation in the voltage is known as the *ripple*.

Now the rectifier will not commence to conduct until the voltage on the secondary of the transformer exceeds the voltage on the condenser. This does not happen at the beginning of the positive half cycle but is delayed until some point towards the top of the wave as shown in Fig. 154. At the point where the transformer voltage overtakes the steady voltage on the condenser a large pulse of current flows through the rectifier causing the condenser voltage to increase. When the input voltage falls below the condenser voltage the rectifier ceases to conduct and the load current is then supplied from the charge stored in the condenser which causes a steady fall in voltage over the idle period as shown.

With a half-wave circuit there is one such pulse of current every cycle, while with a full-wave circuit there are two such pulses, one in each half cycle as illustrated in Fig. 154, where the alternate negative half cycles have been shown reversed (as, in effect, they are by the second rectifier) so that they all appear on one side of the zero line. The voltage on the condenser will continue to rise sharply and then fall away steadily, resulting in a mean voltage output $V$ with a fluctuation above and below the mean value. The total value of this fluctuation $dV$ may be estimated, by assuming that the rectifier is only conducting for a

small fraction of a cycle. Then if $f$ is the number of times per second that the condenser is charged, the loss of charge $dQ$ due to the load current $I$ is $I/f$ (very nearly).

This charge must be replaced by the pulse of current through the rectifier which will raise the voltage on the condenser by the small amount $dV$. But since $V = Q/C$ we can write $dV = dQ/C = I/fC = V/RfC$ where $R$ is the load resistance. The peak value of the ripple is half this value $= V/2RfC$.

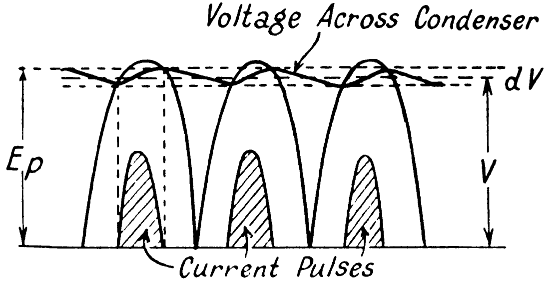



FIG. 154. ILLUSTRATING ACTION OF RESERVOIR CONDENSER

We can re-write this in the form $V = E_p/(1 + 1/2RfC)$. The actual fluctuation will be slightly less than this because the condenser is not discharging for a full half cycle but only about 90 per cent of this time, the balance being occupied by the charging period. The expressions, however, are close enough for practical purposes.

With a half-wave rectifier and a 50-cycle supply $f$ is 50. With a full-wave or bridge rectifier it is 100. With a three-phase rectifier such as we shall discuss later the value may be 150 or 300. This is the frequency of the ripple and it is convenient to ignore the fact that the ripple voltage is not a true sine wave but to consider it as if it were a small sinusoidal voltage of frequency superposed on the steady voltage $V$. We can then calculate suitable smoothing circuits to follow the reservoir condenser $C$ and if these circuits are effective enough to remove the fundamental components they will be still more effective in removing the harmonic components so that this assumption of a sinusoidal ripple is quite justifiable.

## Size and Rating of Reservoir Condenser.

It will be seen from the expressions just developed that the larger the reservoir condenser the nearer does the output voltage approach the peak a.c. The result depends on the product $CR$. If this is small the factor $1/2RfC$ is large and $V$ is considerably less than $E_p$, but as we increase $CR$ by increasing $C$ or $R$ or both, $V$ rises rapidly at first and then progressively more slowly as shown in Fig. 155.

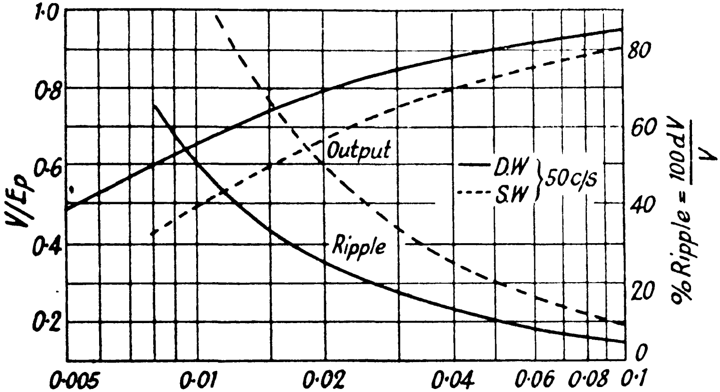



FIG. 155. VARIATION OF OUTPUT AND RIPPLE WITH PRODUCT $CR$

It will be noted that if $R$ is infinite (no load) $V = E_p$. Hence the condenser must be designed to withstand the voltage $E_p$ continuously. It is indeed desirable to allow a small factor of safety for, as we shall see later, leakage inductance in the transformer may provide a resonance which will cause $V$ to exceed $E_p$. This rise is usually small, however, and is often ignored, particularly if a wet type of electrolytic condenser is used for $C$. Such a condenser will break down under excess voltage but will re-seal itself when the voltage falls below the specified surge or peak value.

For normal loads the product $CR$ should be greater than about 0·02. Thus with a load of 5 000 ohms $C$ could be 4 $\mu$F. With smaller currents, such that $R = 50\,000$ ohms, say, a condenser of $\frac{1}{2}$ $\mu$F. would suffice.

## Smoothing Circuits.

Even if $CR = 0{\cdot}02$ there will still be 25 per cent ripple. This, of course, is far too great to be tolerated in a source of h.t. supply to a valve amplifier. We can reduce this by increasing $C$, but this is not economical. For one thing $C$ must not be made too large because the larger the value of $C$ the heavier is the pulse of current through the rectifier and if this current becomes too large it will damage the

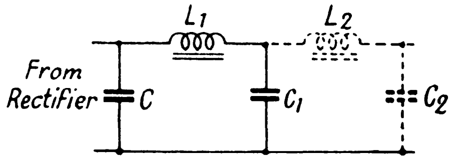



FIG. 156. FILTER CIRCUIT FOR REDUCING THE RIPPLE

rectifier. A value of 8 $\mu$F. is usually the maximum, though 16 $\mu$F. is sometimes used.

Secondly, doubling $C$ will only halve the ripple, whereas if we use the additional capacitance with a series inductance to form a filter, as shown in Fig. 156, we can obtain much more effective smoothing. The voltage applied across the input to the filter is $dV/2$.* This will cause a current $dI = dV/(L_1\omega - 1/C_1\omega)$, neglecting the resistance of the choke. The voltage across the output condenser is $dI/C_1\omega = dV/(L_1C_1\omega^2 - 1)$, so that the ripple is attenuated by the amount indicated by the denominator $L_1C_1\omega^2 - 1$.

If a second section is added this produces a further attenuation of $(L_2C_2\omega^2 - 1)$ making the total attenuation $L_1L_2C_1C_2\omega_4$ very nearly. Such double filters are occasionally required but usually one section suffices.

## Design of Rectifier System.

In practice we desire to arrange our system to deliver a certain d.c. voltage at a certain current. The first step is to decide how much voltage drop we shall experience on our

* It is customary to specify the ripple in terms of its peak value in either direction, i.e. half the maximum variation of voltage.

system. The smoothing choke or chokes will have some d.c. resistance and this must either be measured or estimated and due allowance made for the voltage drop. It is also necessary to allow some additional voltage drop for the loss on the rectifier, and for the resistance of the transformer winding.

The latter effect we shall ignore here, since it can be allowed for by designing the transformer to deliver the voltage required at the specified load current. Further information on this point is given in the next chapter. The rectifier drop depends on the type of rectifier. Except with mercury vapour rectifiers which are considered in the next section we can approximate to the truth by assuming the rectifier to have a mean resistance which can be added to that of the chokes.

With a valve rectifier a value of 250 ohms is representative, while with metal rectifiers $R_r$ may be between 500 and 1 000 ohms. These figures are admittedly indefinite. They are, in any case, approximations because the current through the rectifier is many times the load current so that the actual voltage drop is the product of the instantaneous rectifier current and its resistance in that condition, but it must be remembered that a 2–1 error in the estimation of $R_r$ only makes a difference of a few per cent in the final result since $R_r$ is usually quite small compared with the load resistance $R$.

Because of the difficulty of assessing the rectifier drop it is customary for the valve makers and manufacturers of metal rectifiers to publish data giving the rectified output with a stated value of reservoir condenser in terms of the a.c. input from the transformer for different values of load, so that the designer may suit himself as to what method he adopts. A typical calculation will indicate the procedure. Let us assume we wish to develop a voltage of 300 volts at a current of 60 mA. The load resistance $R$ is thus 5 000 ohms. To make $CR = 0{\cdot}02$ requires $C = 4\ \mu$F. With a full-wave 50 c/s circuit $dV = 1/2RfC$ then becomes 0·25. (This may be read off Fig. 155.)

If the smoothing choke has a resistance of 200 ohms and we assume $R_r$ also equal to 200 ohms the combined choke

and rectifier drop will be $400 \times 0{\cdot}06 = 24$ volts, so that the input voltage required is 324 volts.

Thus $$E_p = V(1 + 1/2RfC)$$
$$= 324\ (1{\cdot}25) = 405.$$

Whence $E_{rms} = 286$ (on load).

The curves of Fig. 157 are for a typical full-wave rectifier. If we use these, we omit the rectifier drop from the above

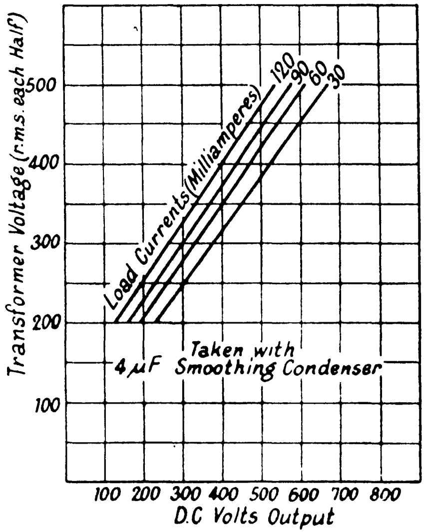



FIG. 157. CURVES FOR A TYPICAL FULL-WAVE RECTIFIER

calculations so that the input voltage required is 3120 From the curves this will be seen to require an r.m.s. input at 60 mA. of 285 volts.

The ripple will be 25 per cent. If we use a 30 H. choke and an 8 $\mu$F. smoothing condenser this will be reduced 94 times, as shown on page 269, leaving 0·265 per cent or $\pm$ 0·8 volts ripple, which is negligible for all normal requirements.

## Voltage Doubler Circuits.

The action of the voltage doubler circuit is a little more

complex than the simple circuits so far considered, because the pulses of current first cause an increase in the voltage on the top condenser and then half a cycle later an increase *in the opposite direction* of the voltage on the bottom condenser. The two condensers swing away from the mid point as shown in Fig. 158. This will be seen to produce a curious shape of ripple and while this has a frequency

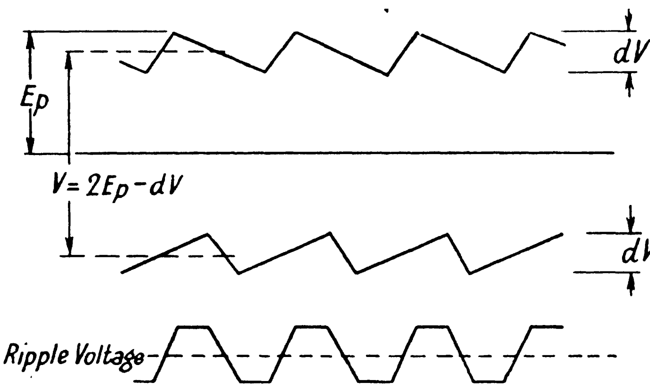



FIG. 158. ILLUSTRATING ACTION OF VOLTAGE DOUBLER CIRCUIT

equal to twice the supply frequency and thus behaves similarly to a full-wave rectifier, the regulation of the circuit is equivalent to that of a half-wave rectifier. This can readily be understood because only half the total capacitance is replenished every half cycle which is equivalent to replenishing the total capacitance every cycle.

It will be clear from the diagram that as a first approximation the output voltage is

$$V = 2E_p - dV$$

$$= 2E_p - V/RfC$$

Whence $$V = 2E_p/(1 + 1/RfC)$$

In the above expressions—

$V$ is the total output voltage,

$E_p$ is the peak transformer voltage,

$C$ is the capacitance of *each* of the voltage doubling condensers (assumed equal), and

$f$ is the *supply frequency* (equals 50 with the normal supply).

The peak ripple is a little less than $dV = V/RfC$. The exact value depends upon the percentage of the time occupied in the charging action but it is satisfactory to take this figure for purposes of calculation. If $C$ is the capacity of *each* condenser, $f$ must be the fundamental frequency as above, despite the fact that the ripple frequency is twice the fundamental. Alternatively we can view the system as being equivalent to a single condenser of capacitance $C/2$ operating at a ripple frequency of 100, which gives the same result.

An example will be of interest. Let us assume a Westinghouse H.T. 16 rectifier required to deliver 300 volts at 60mA. Thus $R = 5\,000$ and we will assume $C = 4\ \mu$F. (each condenser). $RfC$ thus $= 1$, assuming $f = 50$.

Let us assume a rectifier resistance of 500 ohms, causing a drop of 30 volts, so that $V$ should be 330.

$$\text{Then} \qquad 2E_p = V(1 + 1/RfC) = 330 \times 2.$$

$$\text{Whence} \qquad E_{r.m.s.} = 660/2\sqrt{2} = 235.$$

The Westinghouse data book suggests 240 volts which is in good agreement considering the arbitrary assumption of $R_r$.

The ripple will be $V/RfC = V$, so that the condenser voltage will swing from zero to 660 volts. The working voltage of each condenser must thus be at least 350 volts. (The recommended value in the data book is 400.) In passing, it may be noted that this is the worst condition which will arise since even on no load the voltage on the condensers will not exceed $E\sqrt{2} = 340$ volts.

## Choke Input Circuits.

As the load current on the system increases, the peak current taken by the rectifier in charging up the reservoir condenser becomes so large that there is a risk of damage to the rectifier valve. A second disadvantage of the type of

circuit so far discussed is that the process of replenishing the charge in a sudden pulse over a small fraction of the cycle necessarily makes the output voltage very dependent upon the load. If we reduce the load resistance so that more current is drawn from the reservoir condenser the voltage will fall so that we have a steadily dropping voltage characteristic.

The first factor is the more serious because, in order to minimize the loss in the rectifier, which may become more appreciable at high powers, it is usual for large currents to employ rectifiers which contain a small amount of mercury vapour. These valves commence operation as an ordinary thermionic device but the electrons emitted from the cathode encounter molecules of mercury vapour in their transit from cathode to anode. These molecules they ionize and the heavy positive ions drift towards the cathode thus augmenting the current many times.

The process is cumulative and the result is a conducting path of extremely low resistance, the only voltage drop of importance being the ionization potential which is necessary to ionize the molecules, this being of the order of some 15 volts, irrespective of the current.

With such a valve very heavy peaks of current might flow, causing rapid destruction of the valve, but fortunately, both this defect and the inherently bad regulation of the condenser input circuit can be overcome if we can arrange that the current from the rectifier is continuous over the conducting half cycle. This can be achieved by introducing a choke before the first condenser in the system, such an arrangement being known as a *choke input filter*. The arrangement is shown in Fig. 159 for a full-wave rectifier. A choke being a device which tends to maintain a constant current it will clearly not permit sudden pulses of current as shown in Fig. 154, but will actually tend to maintain the current through the rectifier, the choke and the load constant.

Under such conditions the condenser $C_1$ has little effect on the voltage developed, the d.c. voltage being determined by the transformer secondary voltage $E_p$ less the voltage drop in the rectifier and the choke. We have seen that if mercury

vapour rectifiers are used the voltage drop in the rectifier is constant so that we are only left with the voltage drop on the choke and the regulation of the transformer, which involves not only its winding resistance but also the leakage inductance as explained in the next chapter. It is, however, practicable to keep both these voltage drops small so that we are able to provide that the voltage delivered at the output terminals is tolerably constant irrespective of the load current.

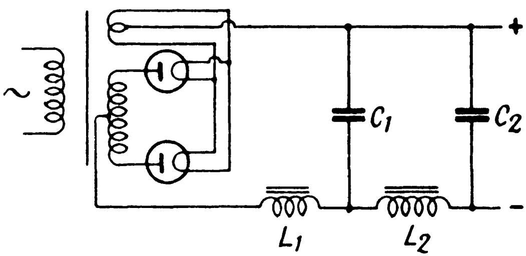



FIG. 159. CHOKE INPUT CIRCUIT

The rise in output voltage from full load current to no load current is called the regulation of the system and it is not difficult to achieve a regulation of 5 per cent only. Actually this good regulation cannot be maintained right down to no load, as we shall see, but over the majority of the current range from about one-tenth full load to full load the change in voltage can be made to be less than 5 per cent.

Fig. 160 shows the form of the current through the rectifier, from which it will be seen that the rectifier current is equal to the load current with a double-frequency ripple superposed and if we increase the drain on the circuit, so that the load current increases, the rectifier current increases in similar proportion.

There is no longer any question of the rectifier not conducting until the a.c. voltage has exceeded a certain amount except just at the very beginning of the cycle. This being so the d.c. voltage output is the mean value of the a.c. from the transformer secondary. We know that the

mean value of a rectified sine wave is $2E_p/\pi$ whereas the r.m.s. value is $E_p/\sqrt{2}$.

Hence the ratio of the r.m.s. voltage on the transformer to the d.c. output is $2E_p/\pi\sqrt{2} = 1 \cdot 11$ (which is, incidentally, the form factor of a sine wave). Conversely the d.c. output is 0·9 times the r.m.s. a.c. input. The actual d.c. at the output terminals, of course, is less than this by

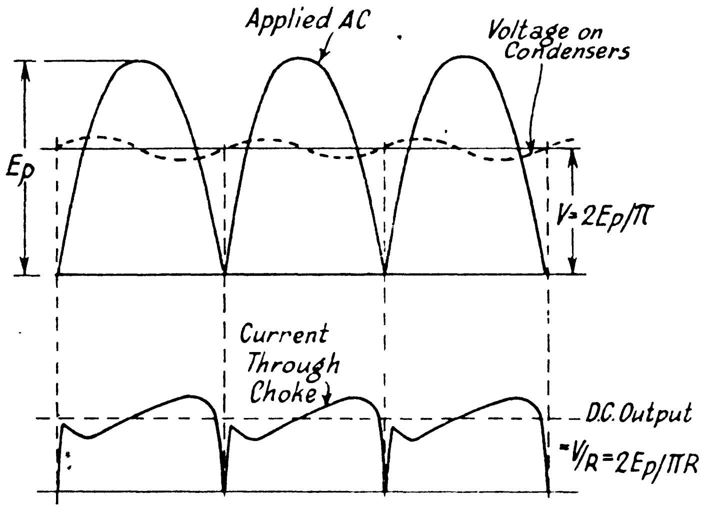



FIG. 160. VOLTAGE AND CURRENT RELATIONSHIP IN CHOKE INPUT CIRCUIT

the amount of the rectifier voltage drop (15 volts) and the drop on the choke due to its d.c. resistance.

## Value of Series Inductance.

The difference in voltage between the applied a.c. from the transformer and the steady d.c. output is absorbed mainly by the choke. The choke will, in fact, carry a ripple current and it will be clear from Fig. 160 that the limiting condition for the maintenance of a continuous current flow through the rectifier is that the peak ripple is less than the steady d.c. current.

Now the form of the ripple voltage delivered by the rectifier is a series of rectified half-sine waves. This may be represented by a mean value $2E_p/\pi$ plus a series of sine waves having frequencies of 2, 4, 6 times the fundamental (50 c/s) frequency. We are only concerned with the first of these—the second harmonic term—since, as we have seen, the effect of the choke will be magnified on the higher harmonics so that if we make it large enough to deal with the lowest frequency ripple it will be more than adequate for the higher frequency terms, which are in any case much smaller.*

The peak amplitude of this second harmonic term is $4E_p/3\pi$ so that the ripple current will be this voltage divided by the impedance of the circuit $Z$. In general the reactance of $C$ will be low compared with the load $R$, so that we can say that $Z = L\omega - 1/C\omega$ ($\omega$ here being the value appropriate to the *double* frequency, i.e. 100 c/s with a 50 cycle full-wave rectifier).

The peak ripple current is thus $4E_p/3\pi\ (L\omega - 1/C\omega)$ and this must be less than the d.c. which is $2E_p/\pi R$ (the mean rectified output divided by the load).

Equating these two we find that $L\omega = 2R/3 + 1/C\omega$. If $C\omega$ is large so that $1/C\omega$ is negligible, this reduces to $L\omega/R = 0{\cdot}67$, a criterion which is often quoted. It may, however, be uneconomical to make $C$ as large as would be necessary to justify this assumption, but it will be clear that if $C$ is reduced, $L$ must be increased.

A very able exposition of the requirements was given by C. R. Dunham (*Journal I.E.E.*, 1943, Vol. 75, p. 278) in which an economical value of $1/C\omega$ of $R/6$ is suggested. This is still small enough to be negligible in comparison with $R$ (remembering that it is in quadrature) and results in a criterion $L\omega/R = 0{\cdot}84$.

## Swinging Chokes.

These values of $L$ are the *minimum* values necessary to maintain a continuous current through the rectifier and the

* The expression for a rectified half-sine wave is

$$e = \frac{2E_p}{\pi}\left(1 - \frac{2}{3}\cos 2\omega t - \frac{2}{15}\cos 4\omega t - \frac{2}{35}\cos 6\,\omega t - \ldots\ldots\right)$$

actual value should be 20 to 30 per cent higher to allow a margin of safety. It will also be noted that the value of $L$ depends on the load. As the load decreases (increasing $R$) the inductance must rise in direct proportion. Chokes can be constructed to comply with this condition over a range of current and are known as *swinging chokes*. They will not maintain their increase below about 1/10th full load, which is the reason for the statement previously made that good regulation cannot be maintained right down to no load, but it is always practicable to place a permanent load across the circuit drawing 10 to 20 per cent of the full load current so that the system always delivers more than the critical current.

It should also be noted that if the load falls below the value for which the choke has been designed so that the current is no longer continuous, the arrangement takes on the characteristics of a shunt-condenser circuit and the voltage rises sharply, reaching, in the limit, the full value $E_p$ at no load.

## Peak Current.

With the criteria quoted above the peak current through the rectifier is limited to twice the mean current, a considerable improvement over the shunt condenser system.

It may, however, be desirable to work with a still lower ratio of peak/mean. Mercury rectifiers are usually rated in terms of peak current, and the permitted ratio of peak/mean may be less than 2, particularly with a small rectifier designed for high efficiency working.

If such is the case the value of $L$ must be increased in due proportion. Thus, if the ripple is halved, the ratio of peak to mean becomes 1·5 instead of 2. Thus, we should achieve the required result by making $L\omega/R = 1{\cdot}68$. It should be noted that increasing $C$ will not suffice because $1/C\omega$ is already small and if it is made infinite it can only reduce the ripple in the ratio $0{\cdot}84/0{\cdot}67 = 1{\cdot}25$.

A final precaution is that the values of $C$ and $L$ should not be such as to resonate with the ripple. This can always be checked but it will be found that with the values suggested

there is no danger of this. If $L$ is under one henry, however, the leakage inductance of the transformer (which is in series with $L$) may increase the value to such a point that resonance occurs. (See next chapter.)

## Smoothing.

It is still necessary to smooth the remaining ripple, for though the choke and condenser constitute a filter the initial ripple may be much larger than with the shunt condenser input circuit. The peak ripple is $4E_p/3\pi$ and since the mean value of the d.c. is $2E_p/\pi$ the ratio of the two is $2/3 = 0{\cdot}67$, which is about three times as great as with a typical shunt condenser circuit.

The circuit is therefore followed, as a rule, by a second filter, the attenuation being calculated as already explained.

A typical design will illustrate all these points. Let us assume our requirements to be a supply of 1 250 volts at a current varying between 0·05 and 0·25 amps. Assuming a double filter with 50 volts drop on each choke and a further 15 volts in the rectifier our input voltage = 1 365. Hence the r.m.s. transformer voltage (on load) must be $1\,365 \times 1{\cdot}11 = 1\,515$.

$$R \text{ is } 1\,250/0{\cdot}25 = 5\,000 \text{ ohms.}$$

Whence $L\omega = 0{\cdot}84 \times 5\,000 = 4\,200$

$$f = 100 \text{ so that } L = 6{\cdot}7 \text{ H. say } 7 \text{ H}$$

This is at full load. At 50 mA. the inductance must be 5 times as great, i.e. 35 H, so that the choke would be specified as 7/35 H at 250/50 mA.

$$C = 6/\omega R = 6/(628 \times 5\,000) = 1{\cdot}93\ \mu\text{F., say } 2\ \mu\text{F.}$$

Initial ripple $= 1\,365 \times 0{\cdot}67 = 915$ volts.

Ripple across first condenser $915/L_1C_1\omega^2$

$$= 915/(7 \times 2 \times 10^6 \times 4\pi^2 \times 10^4 - 1)$$
$$= 915/4{\cdot}5 = 204.$$

If we wish to reduce our ripple to 1 per cent = 12·5 volts we require a further attenuation of 16·5 times which requires an $LC$ product of 42 so that a 7 H. choke and a 6 $\mu$F. condenser would suffice.

**Insulation.**

Because of the high output voltage the chokes would preferably be placed in the negative lead as shown in Fig. 159. This avoids the need to insulate for the d.c. potential but it is still necessary to insulate for the peak ripple in the first choke = 0·67 V. so that for safety the choke should be insulated for a peak working voltage of $V$ and tested at twice this voltage.* If for any reason it is inconvenient to use the chokes in the negative lead the chokes must be insulated for a peak voltage of 2 V.

The first condenser will have to withstand the d.c. voltage $V$ plus the peak ripple across it which we have seen to be 10 and 20 per cent of the total input ripple. Under normal conditions, therefore, it is sufficient to design $C_1$ for a working voltage of 1·5 V. and $C_2$ for about 1·1 V. It should be remembered, however, that if the load on the circuit is removed for any reason the voltage will rise to the full peak value of the a.c. input, on both condensers. Sound design thus requires both condensers to be designed to withstand a peak voltage $E_p$, and to be tested at twice this figure.

The case of the shunt condenser circuit has already been mentioned. The input condenser will have to withstand the full peak value of the applied a.c. = $E_p$ and if a paper condenser is used here the working voltage should be taken as twice the d.c. for safety. The smoothing condenser need normally only be insulated for a voltage $V$, but again if the load is removed it will be subject to the full peak voltage $E_p$, and the same applies to the chokes if these are in the positive lead as is usual. All the components, therefore, should be specified to withstand a peak of 2 V. for safety.

If the first condenser is electrolytic (particularly the wet type) the factor of safety may be dropped. For example typical condensers are specified as 440 V. working, 460 V.

* Condensers are rated in peak voltage and are tested on d.c. An a.c. test on an equivalent voltage (equal to 0·71 times the peak voltage) is not the same because the condenser will draw current. This may cause overheating and does not represent the working condition. When ordering a condenser always specify the conditions under which it will be used. This enables the manufacturer to ensure that the condenser will handle not only the voltage stress but also the current likely to flow.

peak or surge. Such a condenser would be suitable for an a.c. input of 310 volts. It should always be remembered that the a.c. mains are permitted to vary $\pm$ 12 per cent in voltage. 310 volts + 12 per cent gives 492 volts peak which would break down the condenser. If it were of the wet type it would re-seal after each breakdown but continuous use under such conditicns would boil the liquid and quickly render the condenser useless.

The filament windings of the rectifiers must, of course, be insulated to the same extent as the part of the circuit to which they are connected. Thus in Fig. 159 the filament winding must withstand a normal voltage of at least twice the d.c. output voltage.

**Peak Inverse Voltage.**

The rectifiers have to withstand an even greater strain, which is at its maximum during the non-conducting portion of the cycle and is thus known as the *peak inverse voltage.*

Consider Fig. 153 (*a*). On no load the condenser $C$ will charge to $E_p$. During the negative half cycle the point $A$ will be negative so that the transformer voltage will add to the condenser voltage. The peak inverse across the rectifier is thus $2E_p$.

The same applies to each of the rectifiers in Fig. 153 (*b*) and 153 (*d*). In the bridge circuit of Fig. 153 (*c*) the condenser voltage does not add to the transformer voltage so that the inverse peak is simply $E_p$.

**Three-phase Supply.**

Large powers are usually transmitted by a three-phase system. This is an arrangement of three voltages of equal value, displaced from one another 120 degrees in phase. Such a system has many advantages from the power engineer's point of view. For one thing he can obtain three entirely separate circuits with only three wires instead of six. An ideal three-phase circuit in which all the voltages are equal and all the currents are balanced is characterized by the fact that the sum of the three voltages or currents is zero at any instant. This will be clear from the diagram of Fig. 161.

In practice, of course, the loads on the three phases are not exactly balanced so that there is small residual out of

balance current and therefore a fourth or neutral wire is provided to carry this out-of-balance current. It need only be of comparatively light gauge and is usually connected to earth.*

A three-phase supply may be used two ways. The equipment may be designed to operate from each of the three phases to earth as indicated in Fig. 162 (*a*). This is known as a *star connexion*. It may also be designed to operate on the voltages existing between the phases

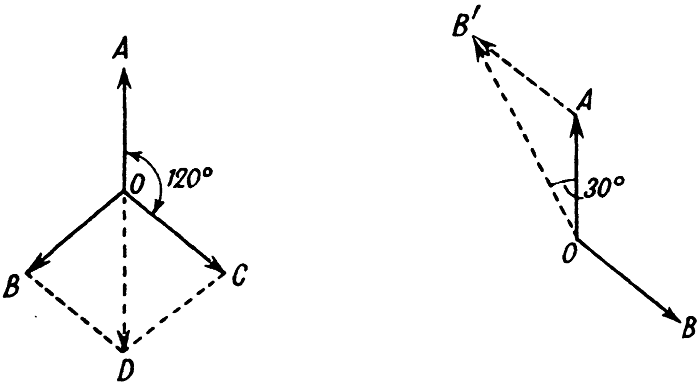



FIG. 161. VECTOR RELATIONSHIPS IN THREE-PHASE SYSTEM

themselves. Now, as is shown in Fig. 161 the vector *difference* between two voltages 120° different in phase is a voltage greater than either, actually $\sqrt{3}$ times as great. It will be clear that if we do this to each pair of phases in turn we shall obtain a similar symmetrical system consisting of three voltages, each $\sqrt{3}$ times the phase voltage, but all adding up to zero as before.

If we design our equipment to operate on the phase

* It is not connected direct to earth for reasons associated with the protection of the generators and transformers at the power station and sub-stations. It is substantially at earth potential but it should not be connected directly to earth in any equipment operating off the three-phase supply. If an isolating transformer is used, however, there is no objection to the connexion of the neutral point on the secondary directly to earth, and this, in fact, is usually done.

voltage, therefore, we still have the three connexions but there is no neutral connexion. This arrangement is known as *delta connexion* and both methods are used in practice, according to which is the more convenient. It is worth noticing in passing that the voltages between the phases are 30 degrees out of phase with the voltages from phase to earth.

Motors of powers above 1 h.p. are nearly always made to

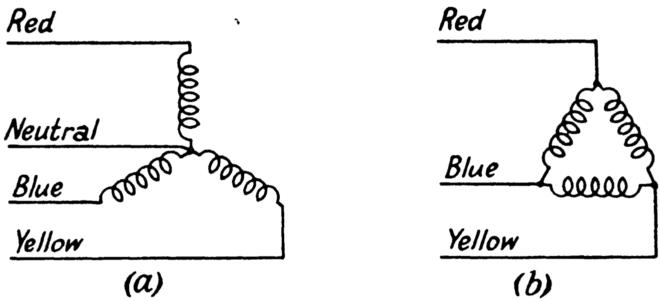



FIG. 162. STAR AND DELTA CONNEXIONS

operate off three-phase supply. This is because if we have three windings spaced 120 degrees apart around the periphery of a stator and we apply three-phase current we obtain the effect of a rotating magnetic field. It will be clear that the voltage will come to a maximum in each of the coils in turn so that the position of maximum magnetic field follows the coils round in sequence and will thus rotate at the same speed as the frequency of the supply, namely 50 times per second. A suitably designed rotating system located in such a rotating field will be dragged round accordingly and such an arrangement provides a very efficient type of motor.

Three-phase transformers can be designed appreciably smaller than for the corresponding power in a single-phase transformer, and generally speaking, for all but the simplest types of equipment, three-phase operation is greatly to be preferred.

The ordinary electric light supply for domestic use is

obtained, merely using one of the phases and the neutral. This gives us a two-wire single phase system. The supply authority arranges its feeder to different districts and even different parts of the same road so as to adjust the loads to be as nearly equal as possible on each of the three phases. It is, therefore, possible to have two houses next door to one another on different phases of the supply, but this does not affect the consumer in any way since each phase acts in an identical manner considered individually.

The fact that the voltage between phases is substantially higher than that between the phase and neutral (earth) is a point which should always be borne in mind. The usual low voltage distributing networks in this country are designed for phase voltages of 230 volts or thereabouts. This is a voltage which is not unduly dangerous except to persons in poor health. The voltage between phases, however, is $230\sqrt{3} = 400$, which is a decidedly dangerous voltage. Due caution should, therefore, be used when dealing with three-phase circuits for this reason and even the insulation of the leads to any transformers or similar equipment may require more attention, for although a comparatively low grade insulation may suffice for the ordinary 230 volt supply the position may be quite different if the insulation has to withstand 400 volts.

## Three-phase Rectifier Circuits.

It will be clear that where three-phase supply is available it is possible to improve the conversion of a.c. to d.c. by using all three phases. If we charge the condenser three times in every period our ripple frequency becomes 150, while if we can arrange a three-phase full-wave arrangement the ripple frequency becomes 300.

Fig. 163 shows a three-phase single-wave rectifier arrangement. The transformer is connected with its secondary in star and a rectifier is connected from the outer of each of the limbs. The cathodes of all three rectifiers are common and are taken to the input circuit. A condenser input circuit may be used for small powers but since three-phase working is usually employed where higher powers are involved one usually finds a choke input circuit. For the

reasons already mentioned the choke is connected in the neutral point so that it is at low potential.

The construction of three-phase transformers will be described in the next chapter. It may be noted, however, that the primary of the transformer may either be a similar star arrangement with each of the windings designed for operation at phase voltage or it may be a delta arrangement

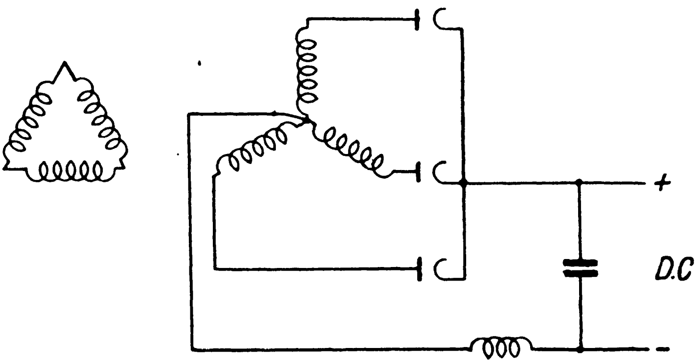



FIG. 163. THREE-PHASE SINGLE-WAVE RECTIFIER CIRCUIT

with the windings designed to operate at the phase-to-phase voltage. For small powers three separate single-phase transformers may be used, but this arrangement is not used for higher powers because with any single-wave transformer the transformer secondary has to carry a d.c. component. The waveform of the secondary is, in fact, the rectified half wave already mentioned, and the d.c. component may produce saturation of the iron, With a three-phase transformer in which the windings are arranged on three limbs of a special type of stamping this saturation is avoided as is explained later.

The design of the filter is modified from that of a single-phase case because of the overlapping of the voltages as shown in Fig. 164. The effective voltage thus never falls to zero so that the ripple is considerably reduced.

Actually the mean value of a three-phase sequence of the type shown in Fig. 164 (*a*) is 0·83 $E_p$ and the peak ripple is

$0 \cdot 3\ E_p$. This ripple has a frequency of three times the fundamental frequency, i.e. 150 c/s for a normal 50 c/s supply.

Using these figures the choke and first condenser may be calculated by the same method as before, using $f = 150$, and the second choke and condenser constituting the ripple filter can be calculated in a similar manner. Owing to the higher frequency of the ripple and smaller initial ripple

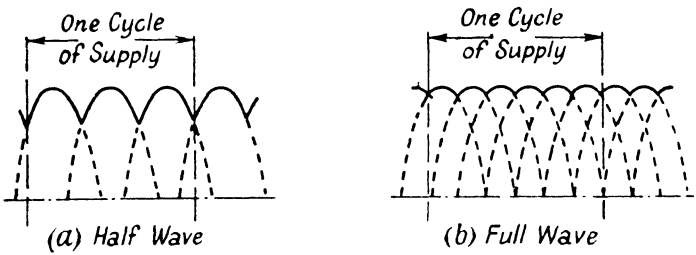



FIG. 164. ILLUSTRATING REDUCED RIPPLE WITH THREE-PHASE SUPPLY

content it is possible that a second filter may not be required. If it is necessary its inductance and capacity will be appreciably smaller than with a single-phase circuit.

Fig. 165 shows a three-phase full-wave rectifier. With this arrangement the successive half cycles of voltage across the load follow one another every 60 degrees, i.e. six times in a complete cycle. The result of this, as will be seen from Fig. 164 (*b*), is that the ripple voltage is extremely small. The voltage 30 degrees on each side of the peak of the wave is $0 \cdot 87$ times the peak value, so that we have only dropped 13 per cent. The total value of the ripple is thus only $6\frac{1}{2}$ per cent and the mean value of the voltage is $0 \cdot 93\ E_p$. Consequently, with such a circuit, smoothing is the least of the problems and it is only necessary to design the choke to maintain a continuous current through the rectifier. The same laws as before apply, but in this case $f$ is 300.

It will be seen that with this circuit two legs of the transformer are always in series, each feeding through its

appropriate rectifier, so that the voltage per leg is only half that which is required in a single-wave circuit. (It must be remembered that we are dealing with peak values so that the phase difference between the voltages on the three legs does not enter into the calculations here.)

Fig. 166 shows yet another form of three-phase full-wave

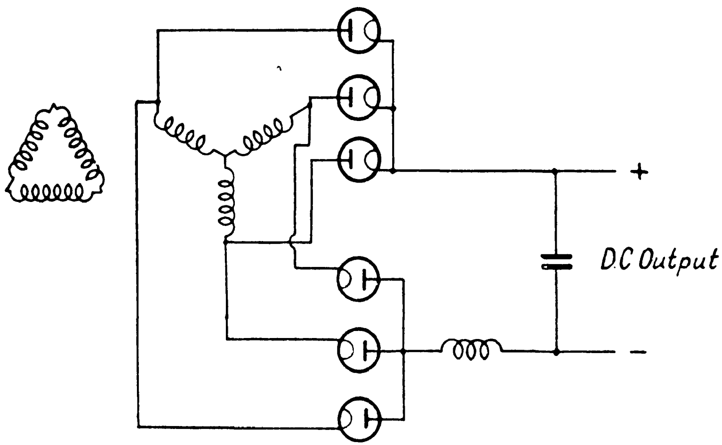



FIG. 165. THREE-PHASE FULL-WAVE RECTIFIER CIRCUIT

rectifier, this being known as the Double Y circuit. Here there are two star connected secondaries so phased that their voltages are, at any instant, in opposition. Thus, the voltage at the point *A* is always equal and opposite to the voltage at point *B*. Each of the limbs of the star is fed to a rectifier but all the rectifiers have common cathodes so that only one filament supply is necessary instead of four separate supplies as in the circuit of Fig. 165.

It is, however, necessary to interpose a reactor or balance coil between the neutral points of each of the stars. This is essential to maintain the continuity of current, and it can, of course, be the ordinary series inductance in the negative lead. If it is arranged in the form shown in Fig. 166 however, it will be observed that the d.c. flows through both halves of the winding in opposition so that the

effect cancels out. We shall see in the next chapter that a choke which has to carry d.c. as well as a.c. has to be larger than one carrying a.c. only because of the saturation of the iron.

## Transformer Current.

So far we have not discussed the current to be supplied by the transformer. The design of transformers is dealt

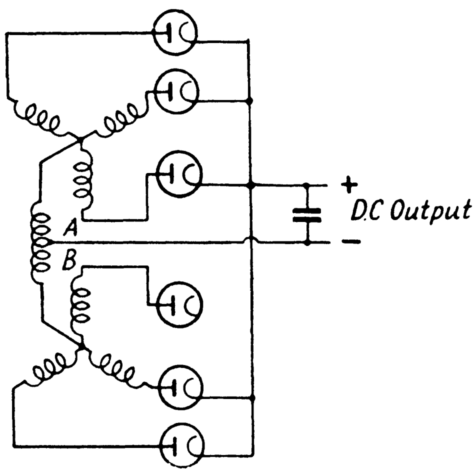



FIG. 166. THREE-PHASE DOUBLE-Y RECTIFIER CIRCUIT

with in the next chapter, but it is relevant to consider here the requirements in a transformer feeding a rectifier.

The choke input type of circuit is the simpler and will be considered first. As we have seen, the current through the rectifier with a single-phase circuit is equal to the d.c. (plus the ripple). This is in the nature of a square-topped wave, as shown in Fig. 167, which is different in its heating effect from a sine wave.

The average value of the secondary current in either limb is $I/2$. We can easily calculate the r.m.s. value; the current squared is $I^2$ for half the time and zero for the remainder, so that the mean squared value is $I^2/2$ and the

root of this (which is the r.m.s. value) is $I/\sqrt{2}$. Hence the effective current in each limb of the secondary winding to $I/\sqrt{2}$ and the transformer should be designed accordingly.

Similar calculations can be made for three-phase circuits. With a three-phase half-wave circuit, for example the current is $I$ for one-third of the time so that the effective r.m.s. current is $I/\sqrt{3}$.

For a bridge or a three-phase full-wave circuit the current

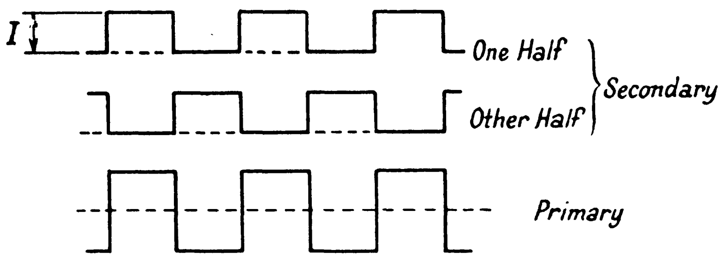



FIG. 167. FORM OF CURRENT THROUGH RECTIFIER

is symmetrical, being a square-topped wave of peak value $I$, alternately positive and negative. The r.m.s. value of such a wave is $I$, as the reader may easily verify for himself.

With shunt condenser input circuits the calculations are not so simple. As we have seen, the current in such a circuit is in the form of large peaks of short duration. Since the heating effect is proportional to the square of the current, such currents cause considerable heating. As a general guide the transformer secondary should be designed to supply (continuously) an r.m.s. current of 1·5 to 2 times the d.c. load current for a full-wave circuit, the ratio becoming larger with smaller d.c. current.* For a half-wave or a voltage doubler circuit the ratio should be between 2 and 3.

The table overleaf summarizes the information.

* It is sometimes suggested that as the transformer secondary in a full-wave circuit only supplies current half the time the above figures may be halved. This is incorrect as the necessary allowance has already been made in the calculations.

# VOLTAGE AND CURRENT RELATIONS IN RECTIFIER CIRCUITS

TYPE OF CIRCUIT

| | Half Wave Condenser Input | Full Wave Condenser Input (Centre-tapped) or bridge) | Voltage Doubler | Full Wave Centre-tapped Choke Input | Bridge Choke Input | Three Phase Half Wave | Three Phase Double Y | Three Phase Full Wave |
|---|---|---|---|---|---|---|---|---|
| **R.M.S. transformer voltage (per limb)** | $V(1 + 1/2RfC)$ | $V(1 + 1/2RfC)$ | $\frac{1}{2}V(1 + 1/RfC)$ | 1·11 V. | 1·11 V. | 0·86 V. | 0·86 V. | 0·43 V. |
| **Peak ripple voltage** | $V/2RfC$ | $V/2RfC$ | $V/RfC$ | 0·67 V. | 0·67 V. | 0·25 V. | 0·06 V. | 0·06 V. |
| **Ripple frequency** | $f$ | $2f$ | $2f$ | $2f$ | $2f$ | $3f$ | $6f$ | $6f$ |
| **R.M.S. transformer current (per limb)** | $2I$ to $3I$ | $1{\cdot}5I$ to $2I$ | $3I$ | $I/\sqrt{2}$ | $I$ | $I/\sqrt{3}$ | $I/\sqrt{3}$ | $I$ |

$V$ is the d.c. voltage output
$I$ is the d.c. load current

# CHAPTER XVII

## TRANSFORMER AND CHOKE DESIGN

A TRANSFORMER consists essentially of an arrangement in which two coils are magnetically coupled to one another. If a varying current is passed through one coil, known as the primary, the changing flux linking with the second coil, called the secondary, induces a voltage therein.

In a perfect transformer all the magnetic flux produced by the primary should link with the secondary. This ideal state of affairs cannot be attained, but for power frequencies and audio frequencies it is possible to approach the ideal by arranging the coils around an iron core. Since the permeability of iron is many times greater than that of air the lines of magnetic force lie mainly within the iron, and if both coils are wound on the same core the greater part of the flux produced by the primary links with the secondary and vice versa.

Before discussing the subject in detail, however, it is clearly desirable to obtain an understanding of the mechanism by which magnetic fields are produced and then to consider briefly the behaviour of iron and magnetic materials generally under the influence of magnetizing forces.

### Magnetizing Force.

Fig. 168 shows a cross-section of a coil with the magnetic field indicated by dotted lines. This field is clearly not uniform but it will be observed all the lines of force pass through the section $AB$ in the centre of the coil. Hence, if we find the field in the centre of the coil we shall have evaluated the total field produced.

Now the magnetic field at a distance $r$ from a small element of wire of length $dx$ carrying a current $I$ is $Idx/r^2$. We consider the coil to be made up of an infinite number of infinitely small elements at gradually increasing distances from the centre (on either side) and we integrate the total.

This results in the expression

$$H = \frac{4\pi}{10} \cdot \frac{In}{l}\left(1 - d^2/l^2\right)$$

where

$I$ is the current in amperes,
$n$ is the number of turns,
$l$ is the length of the coil, and
$d$ is the diameter of the coil.

If the coil is long the correcting term becomes unity, so that

$$\mathbf{H = 1{\cdot}26\ In/l}$$

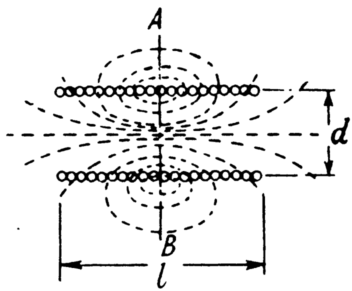


FIG. 168. ILLUSTRATING MAGNETIC FLUX IN A COIL

This is the magnetic field strength at the centre of the coil, also known as the *magnetizing force*. Field strength is defined as the number of lines of force per unit area so that the total number of lines—or the *flux*—is $HA$, where $A$ is the cross-sectional area of the coil. The symbol $\phi$ is usually adopted to represent flux, so that $\phi = HA$.

## Effect of Iron. Magnetic Circuit.

If we insert an iron core in the coil the flux is increased many hundred times. We say this is because iron is more *permeable* than air. It is indeed, so much better that even if the cross-section of the iron is smaller than that of the coil, by far the greater part of the flux inside the coil will be within the iron.

Outside the coil the flux must return through the air as shown dotted in Fig. 169 ($a$), but we shall clearly be still better off if we provide a continuous iron path for the flux as in Fig. 169 ($b$). This is known as a magnetic circuit.

It will also be seen that the iron circuit has the effect of making the lines of magnetic force substantially uniform and parallel to the axis which is equivalent to a lengthening

of the coil. Hence, even if the actual length of the coil is not large compared with its diameter the simplified expression for $H$ still applies. In the case of a complete magnetic circuit, as in Fig. 169 ($b$) the effect is the same whether the coil is wound round part of the iron circuit, as shown, or uniformly distributed over the whole of the iron path. Hence the length $l$ is taken as the mean length of the magnetic path.

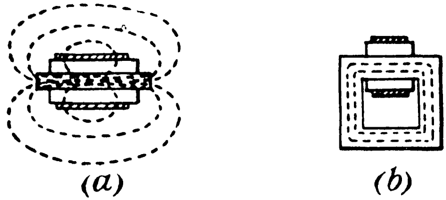


FIG. 169. OPEN AND CLOSED CORES

We allow for the increased magnetic properties of the iron by incorporating a factor $\mu$, representing the *permeability* (which we shall define shortly). Hence we can write $\phi = HA\mu$.

## M.M.F.—Reluctance.

If we write this expression in full we have

$$\phi = HA\mu = 1{\cdot}26In\,A\mu/l = M/S$$

where
$$M = 1{\cdot}26In, \text{ and}$$
$$S = l/A\mu$$

This expression is of similar form to the basic law for electrical circuits $I = E/R$, and although the analogy is not perfect it is often convenient to regard the magnetic circuit as consisting of a magneto-motive force $M$ ($= 1{\cdot}26In$) which forces a flux $\phi$ through a reluctance $l/A\mu$.

On the basis the reluctance is proportional to the length of the magnetic path and inversely proportional to the area and to the permeability. The greater we make $A$ or $\mu$ the less the reluctance and consequently the greater the flux produced, while increasing the length of the magnetic circuit reduces the flux.

### Flux Density—Permeability.

The field strength or *flux density* is the number of lines of force per unit area, so that we can write

$$B = \phi/A = HA\mu/A = \mu H$$

This immediately defines $\mu$, the permeability, as the ratio $B/H$, i.e. the ratio of the flux density with iron present to the magnetizing force $H$ (which is the flux density with an air core).

Iron is not the only magnetic material. Certain other substances, such as nickel, cobalt, chromium and manganese, have permeabilities greater than unity and are said to be *paramagnetic*. Only nickel, however, exhibits the effect to a marked extent and even so is much inferior to iron for which $\mu$ may be several thousand. The chief value of the paramagnetic substances is in the production of alloys with iron, some of which are very highly permeable (e.g. mumetal, an alloy of nickel and iron which can have a permeability exceeding 10 000).

Certain other substances, such as antimony and bismuth, exhibit less permeability than air and are said to be *diamagnetic*. The difference from unity, however, is small and in general substances which are not magnetic, including all insulators, may be taken as having a permeability of unity.

### B-H Curves—Saturation.

The permeability of iron and magnetic materials generally is not constant but varies with the value of $B$, which in turn depends on $H$. One thus needs to know the conditions under which the iron is working, and, to some extent, its past history.

Fig. 170 shows the magnetization or $B$—$H$ curve for soft iron, and also the variation of $\mu$ with $H$. It will be seen that the value of $B$ rises rapidly at first and then much more slowly. This subsequent falling off is known as *saturation*.* It will be seen that $\mu$ rises to a maximum and

* Magnetism may be considered as the result of an orientation of the orbits of the electrons in the atoms each of which forms a small electro magnet. Normally the distribution is random and only a

then falls off rapidly. We shall see later that the effective value of $\mu$ is much below the static value of several thousand shown.

Fig. 171 shows similar $B$—$H$ curve for other materials.

**Hysteresis.**

A further peculiarity is that after iron has been magnetized, the removal of the magnetizing force does not

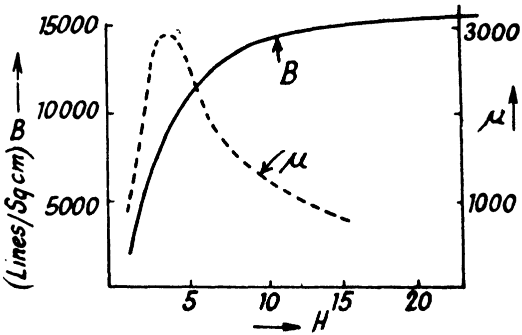



FIG. 170. VARIATION OF $B$ AND $\mu$ WITH $H$

result in complete disappearance of the magnetism. Even if the magnetizing force is only reduced slightly the magnetism will be greater than it was for the same value of $H$ previously. If the magnetizing force is taken over a cycle from a given positive value to a similar negative value and back again, the value of $B$ will follow a curious path as shown in Fig. 172. This is called *hysteresis* and plays an important part in the design of iron circuits.

---

small magnetic effect results, known as residual magnetism or *remanence*. The magnetizing force $H$ causes the orbits to align themselves so that their magnetic effects all tend in the same direction. At first this action is rapid, but after the majority of the orbits have been correctly aligned, increase in $H$ produces little effect and in the limit the only increase produced will be that due to the magnetizing force itself. The material would then be fully saturated. Such a condition is rarely attained in practice, but it is easy to reach a condition where doubling the value of $H$ only produces a 5 per cent increase in $B$.

In a transformer where we apply repeated cycles of a.c. to the coil the iron is taken through this cycle many times a

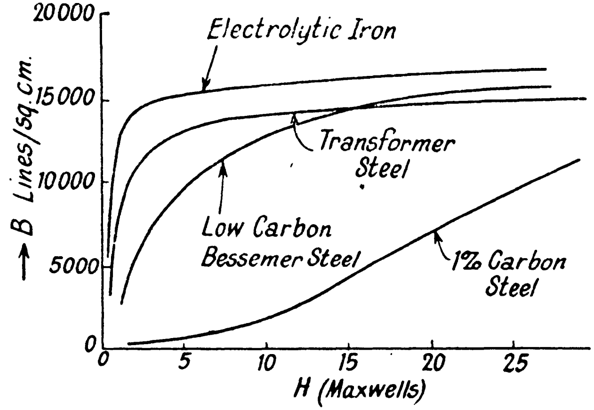



FIG. 171. B—H CURVES FOR VARIOUS MATERIALS

second. It will be clear that the effect of hysteresis is that some of the energy originally applied to the iron is stored in

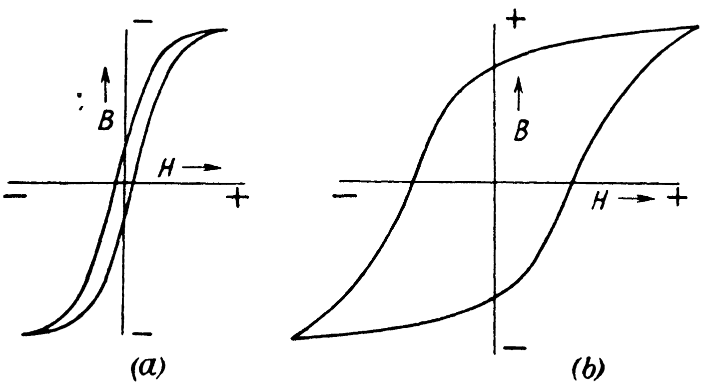



FIG. 172. TYPICAL HYSTERESIS LOOPS FOR (a) SOFT IRON AND (b) MAGNET STEEL

the molecular structure and before we can magnetize the iron in the opposite direction we must dissipate this

energy. There is thus a loss of energy every cycle, which is known as the hysteresis loss.

It is proportional to the maximum flux density to which the iron is magnetized, rather more than in direct proportion but not as much as to the square of $B_{max}$, while it is also directly proportional to the number of cycles per second. The loss is, in fact, of the form $W = kfB_p^{1 \cdot 6}$ where $k$ is a constant depending on the material. It is found that the loss is reduced by incorporating from one to four per cent of silicon in the iron and such material, known by various trade names such as Stalloy, Silcor, etc., is universally used for transformers and chokes.

It is put up in the form of thin sheets as explained later under the heading of *Iron Loss.*

## Remanence.

The amount of magnetism still left when the magnetizing force is removed is known as the *remanence*, while the magnetising force necessary to reduce this to zero is called the *coercive force*. For transformer core material both these should be as low as possible. Occasions arise, however, where we require a high remanence. Carbon steel, for example, has a high remanence and if magnetized retains a high proportion of the energy and becomes a *permanent magnet*. Special alloys such as cobalt steel have been devised for such purposes which have a very high remanence. They require a similarly large coercive force and thus possess a very broad hysteresis loop, which is another way of saying that they retain a high proportion of the energy expended on them in establishing the field in the first place.

## Power Transformers.

After this brief discussion of the magnetic circuit, let us now consider the design of a power transformer which is to receive its supply from the a.c. mains and deliver a secondary voltage which will normally be higher or lower than the input voltage (although sometimes 1 : 1 ratio transformers are used to isolate equipment from the mains). The alternating current flowing through the primary winding will produce an alternating magnetic flux. This flux

linking with the turns of the secondary winding will induce e.m.f.'s therein proportional to the rate of change of flux and the number of turns in the winding. If the flux is varying sinusoidally the secondary voltage will also be sinusoidal so that as long as we ensure that the flux wave is not distorted the transformer will deliver a sine wave output on its secondary.

It is convenient to work from this assumption of an alternating flux as a starting point. This flux will, of course, induce voltages in both primary and secondary windings. Therefore, in the first place if we are to maintain the flux we must apply across the primary a voltage equal and opposite to the back e.m.f. induced in the primary. As long as we continue to do this the flux wave will be maintained and this flux linking with the secondary will induce a voltage proportional to the rate of change of flux.

Let $\phi = \phi_p \sin \omega t$ be the flux, where $\phi_p$ is the peak flux.

Primary back e.m.f. $= e_1 = n_1 d\phi/dt = -n_1\phi_p\omega \cos \omega t$

Secondary e.m.f. $= e_2 = -n_2\phi_p\omega \cos \omega t$.

Thus the ratio of the e.m.f.s on the primary and secondary windings is equal to the turns ratio.

In practice we design the transformer so that $e_1$ is equal to the applied voltage $v$. Now $e_1$ depends on the turns, the frequency of supply $f(= \omega/2\pi)$ and the flux. If the first two factors are fixed the flux will adjust itself until $e_1 = v$. We know that $\phi_p = B_pA$.

Hence we can write $e_1 = B_pAn_1\omega \cos \omega t$.

Hence the r.m.s. primary e.m.f. $E_1 = B_pAn_1\omega/\sqrt{2}$.

If we express all the quantities in the usual units (volts, amps., etc.) we have to introduce a correcting factor of $10^{-8}$. Allowing for this, and writing $2\pi/\sqrt{2}$ in figures, we have

$$\mathbf{E_1 = 4{\cdot}44B_pAn_1f \times 10^{-8}} \text{ volts.}$$

This is the fundamental expression used in transformer design. We choose a suitable value of $B_p$, usually around 10 000, and then adjust the area of core $A$ and the number of turns $n_1$ to make $E_1$ equal to the input voltage required

(e.g. 230). The secondary turns $n_2$ then $= E_2n_1/E_1$, subject to a correction for internal voltage drop as explained later.

## Effect of Secondary Load.

What happens if we connect a load across the secondary? We know, from Lenz's law, that the e.m.f.'s of self and mutual induction are in such a direction as to oppose the change of flux which generates them. Hence if we connect a load across the secondary a current will flow $= E_2/Z_2$ which will be in such a direction as to *reduce* the flux in the core.

But once this happens, the primary e.m.f. $E_1$ is no longer equal to $V$ and hence more current will flow in the primary until the flux is restored to its original value. It is not difficult to see that the relative demagnetizing effect of the secondary will be proportional to the number of secondary turns. If $n_2 = n_1$ the current necessary in the primary to restore the flux is the same as the secondary current. If $n_2$ is greater than $n_1$ then the primary current must be proportionately larger and vice versa, so that $I_1/I_2$ is *inversely* proportional to the turns ratio.

This also follows from consideration of the power in the circuit since $E_1I_1$ must equal $E_2I_2$ neglecting losses. But $E_2 = n_2E_1/n_1$, whence

$$I_1 = I_2n_2/n_1$$

## Magnetizing Current.

There must be some primary current, even when there is no load on the secondary, in order to maintain the flux. This initial current is termed the no-load or *magnetizing* current. It is nearly 90° out of phase with the primary voltage because the primary is almost wholly inductive. There is a small resistive component due to the winding resistance and the losses in the iron, as we shall see later, but in general this is small.

The primary load current, however, is usually in phase with the voltage. Actually, it is of similar phase to the secondary current and if a complete analysis is required vector diagrams should be drawn from which the primary current can be determined as the sum of the magnetizing current and the load current. Simplified vector diagrams,

neglecting losses, are shown in Fig. 173 for resistive, inductive and capacitive secondary loads. The primary voltages and load currents are in opposition to the secondary quantities in each case.

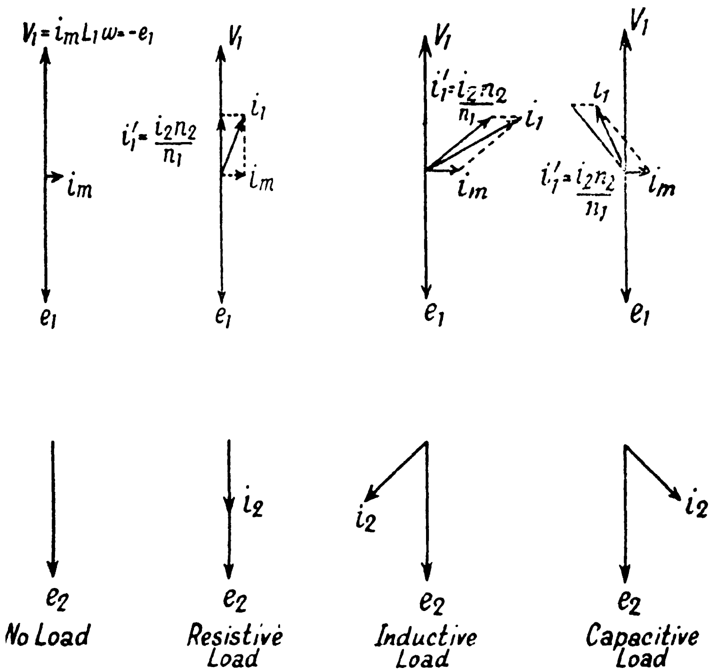



FIG. 173. TYPICAL TRANSFORMER VECTOR DIAGRAMS

An example of a capacitive load is found in a shunt condenser rectifier circuit while inductive loading is experienced with choke-input rectifiers. A heater supply would be mainly resistive.

## Equivalent Circuit of Transformer.

We see, therefore, that the connection of a load across the secondary of a transformer results in a secondary current which is reflected into the primary, the ratio of the

currents being in the inverse ratio of the number of turns. This relation is only part of the story, however, because as soon as we begin to draw a current from the transformer the internal impedance of the device begins to take effect and we find that the output voltage is no longer in simple turns relation to the input voltage. Fig. 174 illustrates the equivalent circuit of a transformer. We commence with the voltage $V$, which is the input voltage. The primary winding, however, has a certain resistance so that the voltage actually applied across the transformer is $V - I_1R_1$.

The secondary e.m.f. will be $E_2 = (V - I_1R_1)n_2/n_1$ and the secondary output voltage will be less than this amount because of the voltage drop on the secondary resistance, so that $V_2 = E_2 - I_2R_2$. It will be seen, therefore, that the turns ratio must be calculated on the internal e.m.f.'s and is actually appreciably greater than $V_2/V_1$ to allow for the internal voltage drop.

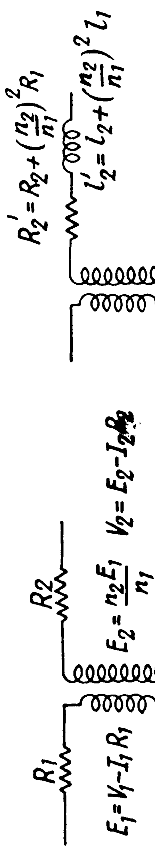



FIG. 174. EQUIVALENT CIRCUIT OF TRANSFORMER

It is often convenient to refer all the quantities to one winding. In the second diagram of Fig. 174 we have assumed all the resistance concentrated in the secondary side and also included an equivalent leakage inductance (see page 302). The equivalent values are obtained by usual laws of coupled circuits (see Chapter IX). Thus,

$$R_2' = R_2 + R_1 n_2^2/n_1^2$$
$$l_2' = l_2 + l_1 n_2^2/n_1^2.$$

With this arrangement $V_1 = E_1$

but $V_2 = E_2 - I_2(R_2' + j\omega l_2')$.

For a more accurate expression see page 304.

## Regulation.

Because of these internal voltage drops, which are clearly dependent upon the current flowing, the voltage output from the secondary of the transformer is not constant but falls steadily as the load increases. The fall in voltage from no load to full load expressed as a fraction of the full-load voltage, is known as the *regulation* of the transformer. This figure varies with the size and quality of the transformer. A small radio set transformer may have a regulation of 15 to 20 per cent. As the size of transformer increases and the quality becomes better the regulation falls to something like 5 per cent, while with a large transformer the figure may be still less.

It is obviously desirable to keep the regulation good because the resistance of the windings not only causes a voltage drop but actually occasions heat losses which are wasteful. If the transformer is a small one it may be that these losses can be tolerated but as the transformer increases in size it is obviously necessary to restrict the losses as otherwise the heat generated becomes considerable.

## Leakage Inductance.

The resistance in the winding is not the only source of voltage drop. We have seen that the object of winding the coils on an iron core is to arrange that all the flux produced by the primary shall link with the secondary and vice versa. This object is not completely maintained and a certain amount of the flux on the primary does not link with the secondary, and the same applies to the secondary.

This is known as the *leakage flux* and the effect will be to introduce additional e.m.f.'s into the circuit. In the primary, for example, in addition to the main e.m.f. generated by the main flux cutting the turns there will be a subsidiary e.m.f. due to the leakage flux. The same applies to the secondary and the simplest way of representing this effect is to introduce into the equivalent circuit a small inductance as shown in Fig. 174. Strictly speaking the leakage flux should be introduced in series with both primary and secondary but it is usually adequate to represent the

whole leakage flux by a leakage inductance in series with either the primary or the secondary only (usually the latter).

The amount of the leakage flux depends on the construction of the transformer. Fig. 175 (*a*) shows a type of transformer which is bad. Here the primary winding is placed on one limb of the core and the secondary winding

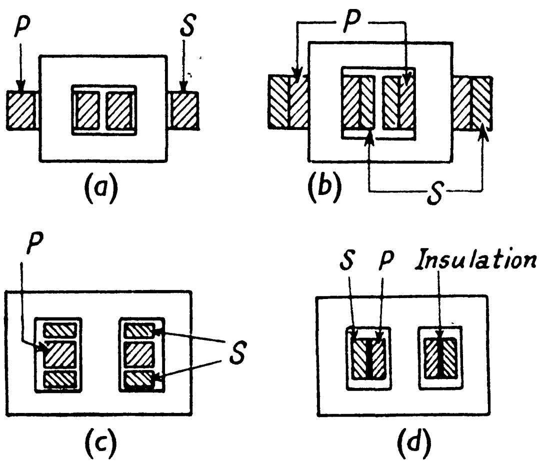



FIG. 175. CORE AND SHELL TYPES OF TRANSFORMER

is placed on the other. It is clear that there must be quite a lot of flux linking the primary which does not link the secondary.

We can improve matters by splitting the winding and putting half the primary on each limb of the core and winding half the secondary over each section of the primary. The sections may be in series or parallel as required. Thus, for a high-voltage transformer we might connect the primaries in parallel and the secondaries in series. Whichever connection is used we must ensure that the directions are correct so that the current at any instant is such as to produce a magnetic field which is in the same direction *around the core.*

This type of transformer is used for medium and high powers and is known as a *core-type* transformer. For small

powers the arrangement of Figs. 175 (*c*) and (*d*) is more usual. Here a three-limb core is adopted, and since the flux has two alternative paths it is clear that the outer limbs, as well as the top and bottom limbs, need only be half the thickness of the centre limb.

This type of construction, known as the *shell type*, has the advantage that only one set of windings is needed, the secondary being wound over the primary as shown in Fig. 175 (*d*). Both these arrangements provide greatly reduced leakage inductance but if still lower leakage is required (as, for example, when feeding a choke-input rectifier which is required to give substantially constant output) still further precautions are required and it is necessary to interleave the windings in the manner shown in Fig. 175 (*c*). This interleaving or sandwiching process clearly minimizes the possibilities of stray flux being able to link with the primary or secondary only and thus reduces the leakage inductance.

Leakage inductance is difficult to calculate and in most design work practical experience is taken as a guide. For a system using concentric windings, however, as in Figs. 175 (*b*) and (*d*), the equivalent leakage inductance (assumed to be all transferred to the secondary as in Fig. 174) is given, to a fair approximation by

$$l = 0{\cdot}0126\,\frac{n^2T}{l_w} \times \left(d + \frac{t_1 + t_2}{3}\right) \text{ microhenries}$$

where $n$ is the secondary turns,

$T$ is the mean turn for the whole of the copper which is as shown on Fig. 176 if the winding fills the "window" space, as it should do if the leakage is to be small. If not, for $B$ write $t_1 + t_2 + d$.

$l_w$ is the length of the shortest winding,

$t_1$ and $t_2$ are the radial thicknesses of the primary and secondary coils, and

$d$ is the distance between the windings.

These parameters are illustrated in Fig. 176. All dimensions are in cm.

It should be remembered that any reactive voltage drop

is in quadrature with the resistive drop so that the total drop is $Iz = I(R + j\omega l)$. This voltage drop is itself out of phase with $V$ by an amount depending on the phase angle of the current $I$, as well as the ratio $\omega l/R$. Hence, it is more accurate to write the voltage drop in the form

$$IR \cos \phi + Il\omega \sin \phi$$

where $\phi$ is the phase difference between the voltage and the current. $R$ is the total equivalent secondary resistance $= R_2 + (n_2/n_1)^2 R_1$

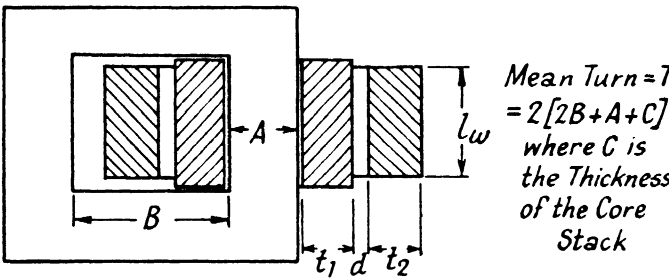



FIG. 176. DIMENSIONS USED IN LEAKAGE FORMULA

**Transformer Losses.**

A practical transformer does not transfer all the primary energy to the secondary, some energy being lost in the process. The losses are of two main forms.

*Copper Loss.* The resistance of the windings causes heating, the loss due to this being $I_1^2R_1 + I_2^2R_2$.

*Iron Loss.* In addition, we have the loss in the iron core due to the varying magnetization. The effect of hysteresis has already been considered while there is a further loss due to the circulating eddy currents set up in the iron which clearly behaves as a short-circuited secondary.

This eddy current loss is quite serious if a solid core is used. In a practical transformer, therefore, the core is built up of thin sheets or strips, insulated from each other by a thin facing of paper or paint on one side. For small transformers the *laminations* are stamped out in the form of T and U pieces or sometimes E and I pieces as shown in Fig. 177. They are assembled with the joints alternately

at each end so that the effect of a substantially solid core is obtained, but built up of insulated thin laminae so that any eddy currents induced can only flow in restricted paths. In addition, the sheets are made of silicon steel which not only reduces hysteresis but has a high electrical resistance so that the eddy currents are still further reduced.

Eddy current loss is proportional to $B_p^2 f^2$ and thus becomes increasingly important as the frequency is raised.

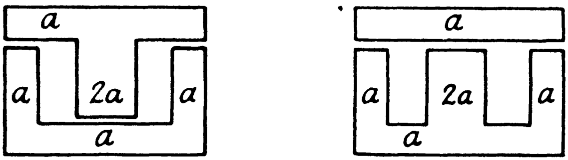



FIG. 177. TYPICAL FORMS OF LAMINATION

It is also dependent on the thickness of the sheet. The usual thickness for 50 c/s is 0·014 inches, though sometimes 0·020 inches is used. For higher frequencies or special circumstances requiring an unusually high value of $B_p$, the stampings may be 0·007 or even 0·005 inches in thickness.

The cost is greatly increased for more stampings are required for a given area of core, while the material itself is more expensive and the stamping of the laminations is more costly owing to the higher precision required in the tools. Otherwise burrs are formed at the edges and these pierce the insulation when the laminations are stacked together, so defeating the object of the use of thin sheets. Even with normal sheets the presence of burrs is to be avoided.

The steel manufacturers market material under various trade names in which the hysteresis and eddy current losses are of the same order at normal frequencies. The total iron loss for typical material is shown in Fig. 178.

There is also a small eddy current loss due to circulating currents in the clamps used to hold the whole structure together but this is usually negligible in small transformers. If there is reason to believe that it is of importance insulated joints must be provided where the bolts come through the clamps and similar precautions taken wherever practicable

to reduce the number of complete conducting paths which might permit eddy currents.

The most efficient design is one in which the iron loss and copper loss are equal but this may not be the most economical design. Iron is cheaper than copper and it may pay, say, to increase the area of the core 50 per cent, with similar increase in iron loss. This would permit the number

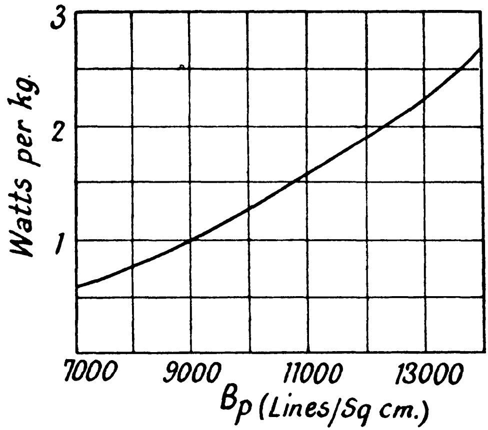



FIG. 178. TOTAL IRON LOSS FOR TYPICAL TRANSFORMER SHEET

of turns to be reduced with saving of copper. In such an instance the copper loss might be only half the iron loss.

The iron loss is present all the time, being a direct result of the magnetization of the iron which does not change with load. It may thus be measured by finding the power taken by the transformer on no load. (This is not $VI_0$, because $I_0$, the magnetizing current, is nearly 90° out of phase.) If it is desired to separate hysteresis and eddy current losses this test must be repeated at two different frequencies.*

* The measured loss is then $W = k_1 f + k_2 f^2$. If this is found for two values of $f$ two equations are obtained which can be solved for $k_1$ and $k_2$. It is important that $B_p$ shall be the same throughout so that the applied voltage must be proportional to frequency. Hence if we use $f_1 = 25$ and $f_2 = 50$, $V_1$ must be twice as great in the 50 c/s test.

The copper loss varies with $I^2$ and is taken at full load. One method is to short circuit the secondary and to apply a reduced voltage to the primary of such a value that the secondary current is the required value. The power taken by the transformer in this condition is the copper loss because $V_1$ will be so small that $B_p$, and hence the iron loss, will be negligible.

These tests require a wattmeter, which may not be available. The losses, however, may be estimated with reasonable accuracy but calculation, basing the iron loss on the curve of Fig. 178 or some similar curve and calculating the copper loss from the resistance of the windings either estimated or measured.

The *efficiency* of a transformer is the ratio of output watts/input watts. For small transformers it is around 80 to 90 per cent, rising to 95 per cent for medium transformers and 99 per cent or even higher for very large transformers. In many cases, in radio practice, the efficiency as such is not as important as ensuring that the losses shall not cause undue temperature rise. As a good practical rule it may be taken that a transformer can safely dissipate $\frac{1}{3}$ watt per square inch of its surface. Thus a transformer having dimensions of $6 \times 5 \times 4$ inches would have a surface area of 148 sq. in., and thus could safely dissipate about 50 watts. Hence it could be used to handle a power of about 500 watts with an efficiency of 90 per cent, since $500/0{\cdot}9 = 555$, equivalent to 55 watts loss.

## Practical Design.

These various factors in transformer design can best be co-related by considering a practical example. Let us take the case of the transformer for the 1 250 volt d.c. supply discussed in the last chapter. The transformer was there shown to be required to deliver 1 515 volts at $0{\cdot}25/\sqrt{2} = 0{\cdot}177$ A. This is per limb, so that the secondary power is $3\,030 \times {\cdot}177 = 536$ watts.

The size of the core is determined by the watts loss to be dissipated. The transformer steel makers supply stampings of various sizes and the transformer designer knows from experience which size of stamping is suitable for a given

wattage (or more accurately volt-ampere) rating. We have just seen that a transformer measuring 6 × 5 × 4 inches would be suitable for about 500 watts and we should, in practice, choose a stamping similar to that shown in Fig. 179, with a core thickness of $2\frac{1}{2}$ in.

The core area would be $0{\cdot}9(1{\cdot}5 \times 2{\cdot}5)6{\cdot}45 = 22$ sq. cm. (The factor 0·9 is to allow for the insulation in the stampings which reduces the effective area of iron.) If we choose

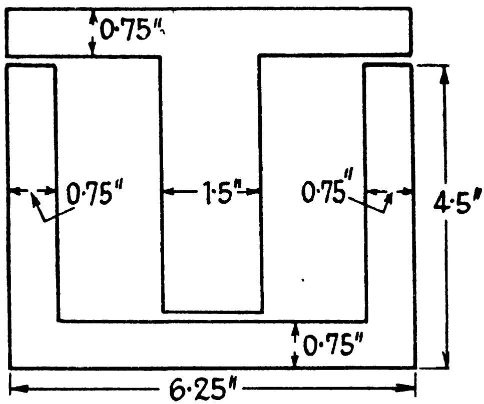



FIG. 179. STAMPING USED IN WORKED EXAMPLE

$B_p = 11\,000$ lines/sq. cm. we can calculate our turns accordingly.* If our primary voltage is to be 230 we have $230 = 4{\cdot}44 \times 50 \times 11\,000 \times 22 \times n_1 \times 10^{-8}$, whence $n_1 = 429$. The secondary turns will be $429 \times 3\,030/230 = 5\,620$.

Now the total window area is $1{\cdot}625 \times 3{\cdot}75 = 6{\cdot}1$ sq. in. But of this about $\frac{1}{8}$ in. will be occupied by the former on which the coil is wound, while we must allow a similar clearance between the outside of the winding and the outer limbs. The winding length will be similarly reduced for there will be the cheeks of the bobbin while, in addition, we shall not wind right to the edge of the bobbin. It is best to allow $\frac{1}{4}$ in. at each end, so that our effective winding area becomes $1{\cdot}375 \times 2{\cdot}25 = 4{\cdot}46$ sq in. We will assume half this to be occupied by each winding.

* A typical value giving roughly equal iron and copper losses as shown on page 311.

The method of winding is usually to wind the turns in layers and to interleave a thin layer of paper between 0·002 and 0·005 in. in thickness between each layer to preserve a uniform winding. For thin wires a 0·002 in. interleave is ample. As a first approximation, assume 90 per cent of the winding area to be available for the wire. Then we have to get 5 620 turns into 0·9 × 2·23 sq in. With the method of winding proposed each wire will occupy a space $= d^2$ so that $d = \sqrt{(0{\cdot}9 \times 2{\cdot}23/5\,620)} = 0{\cdot}0193$.in.

This is the diameter of the wire including its own insulation (silk, enamel, etc.). 27 S.W.G. enamel covered will meet this requirement and since this wire at 1 000 amps./sq. in. will carry 0·21 amps. it is adequate for our purpose.

Now the mean turn (for the whole of the winding space) is 14·5 inches (strictly we should evaluate the mean turn for the primary and secondary separately but the error in taking an average mean turn for both is small). The total length of wire is thus 14·5 × 5 620/36 = 2 260 yds., and from wire tables we find that such a length of 27 S.W.G. wire will have a resistance of 257 ohms.

Since our current is 0·177 amps., our voltage drop will be 45·5 so that $E_2$ must be 3 075·5. Now let us assume that the primary copper loss is the same as the secondary loss, i.e. $I_2^2R_2 = I_1^2R_1$.

Then $$R_1 = R_2I_2^2/I_1^2 = R_2n_1^2/n_2^2 = 1{\cdot}5 \text{ ohms.}$$

The primary current is 0·177 × 5 620/429 = 2·32 amps. so that the primary volt drop = 3·5 volts. Hence $E_1 = 230 - 3{\cdot}5 = 226{\cdot}5$.

We can now correct our turns ratio, which should be 3 075·5/226·5 = 13·59, so that if we leave $n_2 = 5\,620$, $n_1$ becomes 414. Applying the same methods as before we find that this requires $d = 0{\cdot}069$, which corresponds to 16 S.W.G. either enamelled or d.s.c. Taking the same mean turn as before the length of wire = 414 × 14·5/36 = 167 yds. which would have a resistance of 1·3 ohms, thus confirming the assumption above.

It is desirable now to check that the number of turns chosen will, in fact, go into the space. Thus our winding

length we have assumed to be 3·25 in.; 16 S.W.G. d.s.c. wire winds 14·5 turns to the inch so that in each layer we shall obtain 47 turns; 414 turns thus require 9 layers, which occupies 0·62 in. Adding 8 interleaving layers of 3 mil. paper brings this to 0·65 ins. We have $1 \cdot 375/2 = 0 \cdot 69$ in. available. A check on the secondary shows a similar margin so that the design is just correct, remembering that there will need to be at least 0·01 in. of high-grade insulation between the windings because of the high secondary voltage.

The secondary resistance is 257 ohms, while the reflected primary resistance is, by design, of the same order. Hence the copper loss is $514 \times 0 \cdot 177^2 = 16 \cdot 1$ watts, while the iron loss is 9·8 watts (taking the volume in c.c. and using Fig. 178) totalling 25·9 watts. Thus the efficiency is $536/562 = 95 \cdot 3$ per cent.

As a further exercise let us calculate the leakage inductance. $T = 14 \cdot 5 \times 2 \cdot 54 = 36 \cdot 8$ cm. $t_1 = t_2 = 0 \cdot 65 \times 2 \cdot 54 = 1 \cdot 65$ cm. $l_w = 3 \cdot 25 \times 2 \cdot 54 = 8 \cdot 25$ cm. Now we are using the transformer in a full-wave circuit so that only half the secondary is in use at any instant. But the outer half of the secondary is separated from the primary by the distance occupied by the inner secondary.

We can either estimate the leakage for the whole winding and take one quarter of this or assume $d = t_1/2$ and only consider half the turns. In the first case $d = 0 \cdot 01$ plus half the radius of the wire—say 0·06 in = 0·15 cm. total. The formula on page 304 then gives 2·9 H., or 0·71 H. per half. The alternative calculation gives 0·86 H. Let us take 0·8 H. as an average.

The reactance at the ripple frequency is $0 \cdot 8 \times 628 = 502$ ohms. The resistance of half the secondary is 128·5 ohms and if we again allow an equal amount for the reflected primary resistance we have $R_{eff} = 257$ ohms. The equivalent secondary impedance is thus $(257 + j502) = 563$ ohms. The current is lagging by nearly 90° so that, using the expression on page 305 we obtain a volt drop of 89 volts—a regulation of 6·5 per cent, which is about twice what we should expect on the score of resistance alone. (As there will be a further volt drop on the chokes the overall

regulation will be nearer 15 per cent—still quite a good figure for a d.c. supply system, but not as good as can be attained.)

This design has been worked out in some detail to illustrate the method. For many small transformers the winding resistance is relatively higher, giving efficiencies between 80 and 90 per cent, while the leakage inductance calculation is only occasionally required.

## Audio Frequency Transformers.

The basic principles outlined so far apply to other types of transformer. There is a tendency to-day in aircraft to generate much higher frequencies than 50 c/s, frequencies of 400 and 700 c/s being common. The same methods apply and it will at once be seen that the size of the transformer is much reduced by such technique, while the smoothing problem in any rectifier system is also greatly simplified owing to the much higher ripple frequencies.

The iron loss is considerably heavier and the stampings are made thinner in consequence, though with care even the 0·014 in. stampings can be used if some increase in iron loss is permitted.

For audio frequency work the same principles should be used. Often it is possible to work with lower values of $B_p$ in intervalve and output transformers. Leakage inductance is of great importance as the frequency rises and interleaving of the windings are practically essential. (This applies also to 500 c/s power transformers.)

Self capacitance becomes troublesome above 5 000 c/s. It is customary to design the transformer so that the leakage inductance resonates with the self capacitance to maintain the output voltage at the upper frequencies as explained on page 119.

When using some of the special alloys such as mumetal the calculations become complex because the permeability varies with frequency, while the losses also obey a complex law. The design of such transformers thus requires a special technique which cannot be considered in detail here.

## Three-phase Transformers.

Three-phase transformers may be made up as three separate single-phase transformers connected in star or delta as required. Often, however, the three phases are accommodated on a special core having three limbs of *equal* cross section.

Such a core is illustrated in Fig. 180. Let us assume that the red phase is at its maximum voltage. This will produce full flux in the first limb and this flux will complete its

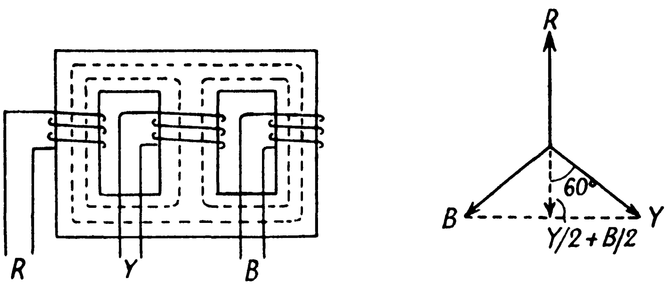



FIG. 180. THREE-PHASE CORE

circuit by passing half through each of the other two limbs The back e.m.f. induced in these limbs, therefore, is half the value of the red phase voltage in opposition. It will be seen from the vector diagram that the component of the yellow phase in the vertical direction is $Y/2$ in opposition and similarly for the blue phase so that the flux in the second and third limbs is exactly what is required for the conditions which are obtained and are in conformity with the fundamental law of three-phase circuits—the sum of the flux in each of the three limbs is always zero.

The area of each limb is therefore designed in accordance with the same principles as for a single-phase transformer and the windings calculated accordingly.

It is worth noting that this same property is of value in a transformer which has to supply a three-phase half-wave rectifier. With any half-wave rectifier the transformer has to carry not only the a.c. but also a steady d.c. current,

This affects the core magnetization as explained earlier, but in a three-phase three-limb transformer such as this the d.c. components cancel out.

As was explained in the last chapter the primaries and secondaries may be connected either in star or delta according to requirements. Any more detailed discussion of three-phase transformers, however, is beyond the scope of this brief review and the reader should refer to standard works such as *Electrical Technology*, by Cotton (Pitman), or the *J. & P. Transformer Book* (Johnson & Phillips).

## Iron-cored Inductances.

We may next turn to a consideration of iron-cored inductances or *chokes*. We have seen that the flux in a circuit is $HA\mu$, and since inductance is, by definition, the linkage per unit current then, neglecting leakage, so that the linkage is merely $\phi n$ we can write $L = HAn\mu/I$. Substituting for $H$ and incorporating the same factor $10^{-8}$ as was used in the transformer expression (see page 298) we have $L = 1{\cdot}26An^2\mu/l \times 10^{-8}$ henries.

Thus the inductance is proportional to the square of the turns, the area of the core and inversely proportional to the length of the core. It is, however, directly dependent on $\mu$, which we have seen to be a very variable quantity depending entirely on the degree of magnetization. Thus the inductance of a transformer primary is varying from instant to instant and it would be difficult to give it any value that had meaning.

A choke is usually required to carry a steady d.c. (though this is not always the case). It thus has a steady magnetization with a ripple superposed. The permeability of the iron will vary with the amount of steady magnetization rather in the manner shown in Fig. 170 so that (except possibly for very low values of $H$) the inductance falls off steadily as the d.c. increases. This effect is known as *saturation* and is illustrated in Fig. 181. Sometimes this effect is of use (as in a "swinging" choke for use with a choke-input rectifier system) but as a rule it is preferable to avoid this dependence of inductance upon the current. This is the more so since the permeability is affected not only by

the d.c. but also by the a.c. component, the inductance tending to increase as the ripple becomes larger.

**Incremental Permeability.**

The effective permeability, moreover, is not the static value for the particular value of *H*. It is found that if an a.c. component of *H* is superposed on a steady value the

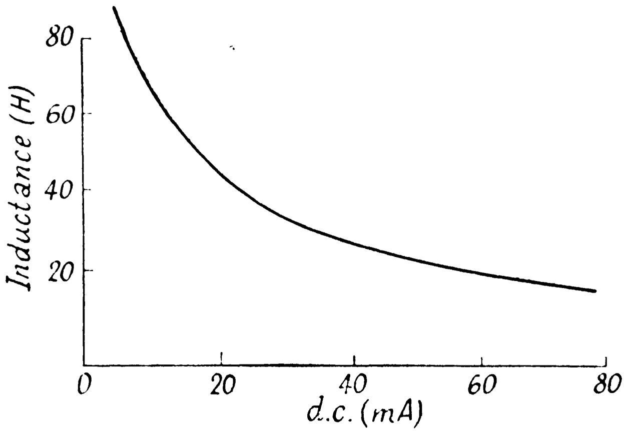



FIG. 181. ILLUSTRATING VARIATION OF INDUCTANCE WITH D.C. CURRENT

iron goes through a small hysteresis loop as shown in Fig. 182, and the effective permeability of the iron is clearly dependent on the average slope of this loop (the dotted line in Fig. 182).

It will be seen that this *incremental permeability*, as it is called, depends upon both $H_{dc}$ and $H_{ac}$. In general

(i) $\mu_i$ decreases as $H_{dc}$ increases;

(ii) $\mu_i$ increases as $H_{ac}$ increases.

The curves of Fig. 183 show values of $\mu_i$ for typical silicon steel sheet. Similar figures are published by the makers for other materials including the special high-permeability alloys such as radio metal and mumetal.

It may be noted, in passing, that a similar effect occurs in a transformer feeding a half-wave rectifier. The winding

in such a circuit has to carry a d.c. component as well as an a.c. component of large magnitude (at least equal to the d.c. and often greater). Here, however, the a.c. component is so large that the principal effect is merely a displacement of the whole loop as shown in Fig. 184. The resulting iron loss is not greatly different from the value without d.c. but the flux wave is no longer even approximately sinusoidal. Hence the primary current is

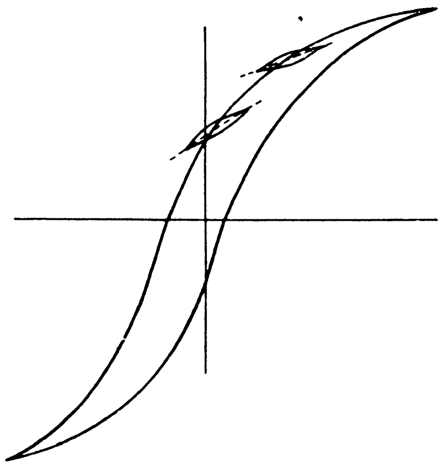


FIG. 182. ILLUSTRATING INCREMENTAL PERMEABILITY

distorted and the secondary waveform becomes peaky in character and contains considerable second harmonic which will result in a greater ripple than that suggested by the simplified calculations in Chapter XVI.

## Use of Air Gap.

The effects of varying permeability may be offset to a considerable extent by including in the iron circuit a small air gap. Even a small gap (of 0·01 to 0·05 in in the average choke) has a reluctance many times the iron so that the performance is determined mainly by the air gap which is of constant permeability.

The flux is reduced because of the increased reluctance.

but the permeability increases to an extent which may more than compensate for this, so that the inductance with an air gap is often greater, for a given d.c. through the winding, rhan with no gap.

Fig. 185 shows a magnetic circuit having an iron path $l_i$ and an air gap $l_a$. The reluctances of these paths are $l_i/A\mu$ and $l_a/A$ respectively. The m.m.f. $M = \phi S$, so that

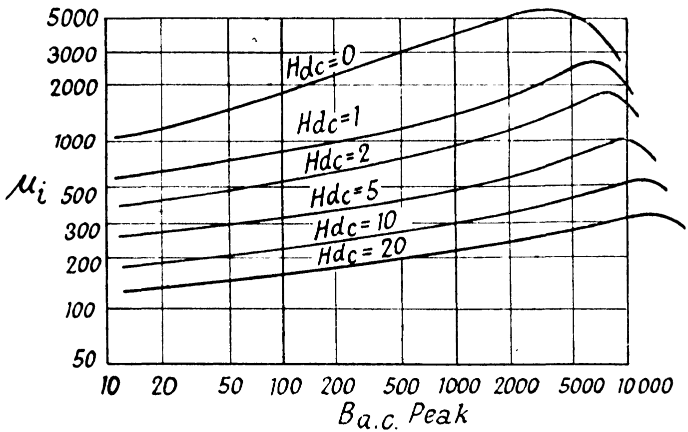



FIG. 183. TYPICAL VALUES OF INCREMENTAL PERMEABILITY

$$M_1 = \phi l_i/A\mu \text{ and } M_2 = \phi l_a/A$$

$$\therefore \text{ Total} \qquad MMF = \frac{\phi}{A}(l_i/\mu + l_a) = \frac{\phi}{A\mu}(l_i + \mu l_a)$$

Now if there is no air gap

$$M = \frac{\phi}{A\mu} . l_i,$$

so that the effect of the air gap may be taken as increasing the length of the magnetic path by $\mu l_a$ (assuming $A$ to be the same, which is usually valid).

It will be clear that if the number of turns and the current through the coil remain unchanged the flux will be reduced in the ratio $l_i/(l_i + \mu l_a)$. If we wish to maintain the flux at its original value the ampere turns must be

correspondingly increased. In the case of a choke we do not wish to do this—in fact, the object of introducing the gap is to reduce the flux—but with some air-gap circuits, such as the field of a loud speaker "pot," we do wish to maintain the flux, and this method of calculation applies equally to either usage.

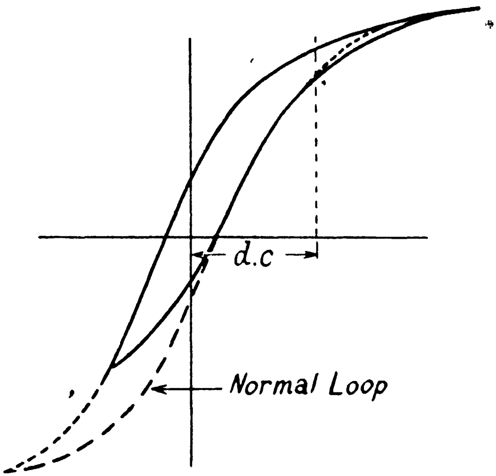



FIG. 184. DISPLACEMENT OF HYSTERESIS LOOP IN SINGLE-WAVE RECTIFIER TRANSFORMER

**Optimum Air Gap.**

Now let the ratio $l_i/l_a = p$. Then the effective length of the magnetic circuit is $l_i(1 + \mu/p)$. We can thus write for the inductance of our choke

$$L = \frac{HAn\mu}{l_i(1 + \mu/p)} = \frac{1{\cdot}26An^2}{l_i} \cdot \frac{\mu p}{p + \mu}$$

The first term is constant, while the second depends on $p$. But as we reduce $p$ (increasing air gap) $\mu$ increases and vice versa. In fact, over a wide range of practical values we are not greatly in error by assuming the product $\mu p$ to constant.

On this assumption we may write $\mu = k/p$. The second term in the expression for inductance then becomes

$1/(p + k/p)$. If we plot this expression in terms of $p$ we find that it rises to a maximum and then begins to fall off again, and we find (either graphically or by calculus) that this maximum occurs when $p = \mu$.

This is another way of saying that the reluctances of the air and iron paths are equal which the reader will recognize as a condition often found in electrical practice (e.g. maximum power when internal and external resistances are equal). The value of $\mu$, of course, is the incremental permeability.

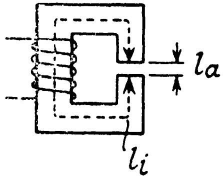



FIG. 185
GAPPED CORE

The best results therefore are obtained with a gap ratio approximately to $\mu$. With a choke having $l_i = 20$ cm. and $\mu = 400$ this would give a gap of

$$p = \mu = 400 = l_i/l_a = 20/l_a,$$

whence $l_a = 0{\cdot}5$ mm. With such a gap the inductance, for a given $H_{dc}$, would not differ greatly from the value at the same $H_{dc}$ without a gap. In most cases it will be found to exceed the ungapped figure and, of course, the inductance with the gap is much more constant despite variations both of the d.c. and the a.c. ripple current through the choke.

It is not practicable to express this improvement in simple terms because the various factors are so interdependent. The simple result just obtained for the optimum gap only arises from the assumption that $\mu p = k$ which is by no means rigidly true. The practical designer, however, is able to make his calculations with comparative ease by the use of tables or curves drawn up partly from more extended theory and partly from empirical data. A very good example of this type of information is to be found in *Radio Data Charts*, by R. T. Beatty (*Wireless World*), which contains a series of *abacs* including suitable transformer and choke data.

# APPENDIX

## ENGINEERING MATHEMATICS

THROUGHOUT this book every attempt has been made to exclude any but the simplest mathematics. The reader may, however, occasionally experience difficulty in working out the details of particular circuits owing to the evolution of unwieldy expressions which are difficult to solve.

The mathematician can often short-circuit much of the working out, and this appendix deals with some of the simpler mathematics which can be used for the purpose. Obviously no proof of the statements can be given, nor is it suggested that the reader should memorize the expressions. A knowledge of the possible methods, however, may prove helpful.

### Rationalization.

Electrical circuit theory is often simplified by the use of the $j$ notation. As explained in Volume I, such expressions can be treated algebraically, but a difficulty arises when fractions are obtained having $j$ terms in both numerator and denominator. For example, the expression $(a + jb)/(c + jd)$ represents some "complex" quantity, i.e. a quantity composed of two vectors at right angles, but it is difficult to see the values of the normal and reactive components.

We therefore *rationalize* the expression by converting the denominator into a term containing no reactive term. If $(c + jd)$ is multiplied by $(c - jd)$ we obtain $c^2 - j^2d^2 = c^2 + d^2$, which is of the required form. We must, of course, multiply the numerator by the same amount, so that

$$\frac{a + jb}{c + jd} = \frac{(a + jb)(c - jd)}{c^2 + d^2}$$

$$= \frac{(ac - j^2 \, . \, bd) + j(bc - ad)}{c^2 + d^2}$$

$$= \frac{ac + bd}{c^2 + d^2} + j\frac{bc - ad}{c^2 + d^2}.$$

We have thus separated the two vectors as required.

This method is very useful. It should be noted that although $c + jd$ is the vector form of $\sqrt{(c^2 + d^2)}$ it is not correct to assume that $(c + jd)^2 = c^2 + d^2$. The term $c + jd$ is a vector expression of which $\sqrt{(c^2 + d^2)}$ is the scalar or numerical value only.

## Network Theory.

It is often desired to know the currents or voltages in various parts of a network. The links may be simply resistive or may be reactive. The usual method is to use Kirchhoff's Law that the sum of the currents meeting at any point is zero, coupled with the use of *cyclic currents*.

For example, considering the simple circuit of Fig. (i), we can assume two separate currents $i_1$ and $i_2$ in each loop. Then we can write

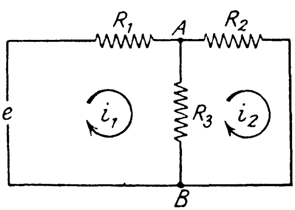



FIG. (i). SIMPLE T NETWORK

$$e = i_1R_1 + (i_1 - i_2)R_3 \quad \text{since } i_1 \text{ and } i_2 \text{ flow through}$$
$$0 = i_2R_2 + (i_2 - i_1)R_3 \quad R_3 \text{ in opposite directions.}$$

This supplies two simultaneous equations which can, after simplification, be solved in the usual way.

## Principle of Superposition.

It is convenient to remember that the various sources of e.m.f. in a network act independently of each other and hence the action of each may be considered separately. This is known as the principle of superposition, the full statement being as follows.

The current at any point (or the voltage between any two points) in a linear network due to the simultaneous action of a number of e.m.f.s distributed throughout the network is the sum of the currents (or voltages) which would exist at these points if the e.m.f.s were acting separately.

The remaining sources are therefore considered as delivering no e.m.f. *but any internal impedance which they possess must still be considered as part of the circuit.*

A linear network is one in which the current in each element is proportional to the voltage across it. The theorem does not apply to networks containing non-linear elements such as an iron-cored inductance or a rectifier or unilateral elements such as valves.

**Reciprocity Theorem.**

Another network theorem, known as the reciprocity theorem, states that if any source of e.m.f. located at a given point produces a certain current at some other point, then the same source of e.m.f. acting at the second point will produce a similar current at the first point.

Thus for either position of e.m.f. and current their ratio, which is called the *transfer impedance*, is the same and the circuit need only be analysed in one direction.

As before, the theorem is limited to networks containing linear bilateral impedances, i.e. impedances which pass current equally in both directions in proportional fashion.

**Compensation Theorem.**

When it is desired to assess the effect of a change in some part of a network (a frequent requirement in circuit calculations) some simplification may be obtained by use of the compensation theorem.

This states that if the impedance of any branch of a network is changed by an amount $\Delta Z$ the current change produced thereby *at any point in the network* is equal to the current which would be produced at that point by an e.m.f. $-I\Delta Z$ in series with the modified branch, where $I$ is the current which originally flowed in that branch.

The theorem is still valid if we remove the impedance completely. In other words, any impedance $Z$ in a network may be replaced by a generator of zero internal impedance, delivering an e.m.f. $IZ$. ($\Delta Z$ here is $= -Z$.)

**Equivalent Networks.**

It is often required to replace one form of network by another of simpler or more convenient form. An example of

this is the frequency bridge of Fig. 75, where the conventional four-terminal arrangement is replaced by a three-terminal network. Such transformations can be made by arranging that the impedances of the two networks are the same under the following three conditions—

1. Looking from the input end with the output open-circuited.
2. Looking from the input end with the output short-circuited.
3. Looking from the output end with the input open-circuited.

Let us apply this to the network of Fig. (ii) (*a*) and derive the equivalent T network of Fig. (ii) (*b*). The original

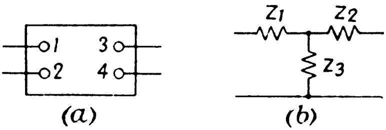



FIG. (ii). FOUR-TERMINAL NETWORK AND EQUIVALENT T NETWORK

network can be any four-terminal network, and the first step is to calculate the three impedances mentioned above, which we will call $Z_{oc}$, $Z_{sc}$ and $Z'_{oc}$.

The same three impedances for the equivalent T network may be written down in terms of $Z_1$, $Z_2$ and $Z_3$ so that, equating the two sets of impedances, we have

$$Z_{oc} = Z_1 + Z_3;$$
$$Z_{sc} = Z_1 + Z_2Z_3/(Z_2 + Z_3);$$
$$Z'_{oc} = Z_2 + Z_3.$$

Re-arranging these equations we obtain values of $Z_1$, $Z_2$ and $Z_3$ as under—

$$Z_3 = \sqrt{[(Z_{oc} - Z_{sc})Z'_{oc}]};$$
$$Z_1 = Z_{oc} - Z_3;$$
$$Z_2 = Z'_{oc} - Z_3.$$

The reader will find it a convenient exercise to apply these rules to the circuit of Fig. 75 (*a*) when he will obtain the equivalent T of Fig. 75 (*b*).

Sometimes one requires an equivalent $\pi$ network. This

may be derived by applying the same basic rules, but the resulting expressions are more elaborate than for the equivalent T. It is often simpler to derive the equivalent T first and then deduce the equivalent $\pi$ by use of the star-mesh transformation.

## Star-Mesh Transformation.

This is a transformation by which a network consisting of a number of impedances meeting at a point, as at Fig. (iii) (*a*),

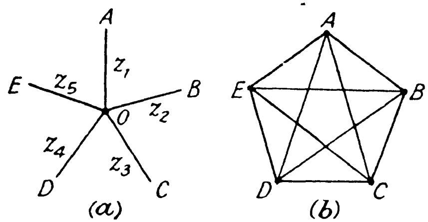



FIG. (iii). STAR NETWORK WITH EQUIVALENT MESH

can be replaced by equivalent impedances joining all the points, calculated as follows.

Let the original star impedances be $Z_1$, $Z_2$, $Z_3$, etc.

Then the mesh impedance between any two points is the product of the star impedances at those points multiplied by, $[(1/Z_1) + (1/Z_2) + (1/Z_3) + \ldots] = \Sigma(1/Z.)$

For example, in Fig. (iii) (*b*), the mesh impedance $AB = Z_1Z_2\ [\Sigma(1/Z)\ ]$. Impedance $AC = Z_1Z_3\ [\Sigma(1/Z)\ ]$ and the general impedance $= Z_mZ_n\ [\Sigma(1/Z)\ ]$.

It may be more convenient to use admittance (= 1/impedance) because with two conductors in parallel the admittances are simply added together whereas the impedance has to be evaluated by the usual expression $1/Z = 1/Z_1 + 1/Z_2$.

If the admittances of the various branches are $A_1$, $A_2$, $A_3$, etc. (= $1/Z_1$, $1/Z_2$, $1/Z_3$, etc.), the mesh admittance is simply $A_mA_n/\Sigma(A)$.

Let us apply this to transform the T network of Fig. (iv)(*a*) into the equivalent $\pi$ network of Fig. (iv)(*b*).

$$Z_a = Z_1Z_3[1/Z_1 + 1/Z_2 + 1/Z_3]$$

$$= Z_1 Z_3 \left[ \frac{Z_1Z_2 + Z_2Z_3 + Z_3Z_1}{Z_1Z_2Z_3} \right]$$

$$= \frac{Z_1Z_2 + Z_2Z_3 + Z_3Z_1}{Z_2} = Z_1 + Z_3 + Z_3Z_1/Z_2.$$

Similarly

$$Z_b = Z_1 + Z_2 + Z_1Z_2/Z_3$$

and $\quad Z_c = Z_2 + Z_3 + Z_2Z_3/Z_1.$

**Thevenin's Theorem.**

There is one other transformation which is often of

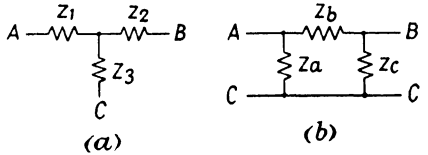



FIG. (iv). T NETWORK AND EQUIVALENT $\pi$ NETWORK

considerable use. This is Thevenin's Theorem, which states quite simply—

The current in any branch of a network is that which would result from the application of an e.m.f. $E$ equal to the e.m.f. available across the specified points *when the branch is removed*, operating on the impedance of the particular branch plus the impedance of the remainder of the network between the points in question.

The procedure will be clear if we consider a simple example. Suppose we wish to know the current in $R_3$ of Fig. (i).

If $R_3$ is removed, the voltage across the points $AB$ will be $E = eR_2/(R_1 + R_2)$. The impedance of the network at these points, still with $R_3$ removed, will be that of $R_1$ and $R_2$ in parallel $= R_1R_2/(R_1 + R_2) = Z$.

Then by Thevenin's Theorem the current in $R_3$ will be $E/(R_3 + Z)$

$$= \frac{eR_2/(R_1 + R_2)}{R_3 + R_1R_2/(R_1 + R_2)}$$

$$= \frac{eR_2}{R_1R_2 + R_2R_3 + R_3R_1}$$

This is a speedy solution to a problem which, though simple, would still require appreciably more calculation by ordinary methods, and in more complex networks the saving can be even more pronounced.

**Determinants.**

We may conclude this brief discussion with some reference to determinants. These are expressions of the form $A_1B_2 - B_1A_2$ which often occur in the solution of equations. In determinant notation this is written, $\begin{vmatrix} A_1 B_1 \\ A_2 B_2 \end{vmatrix}$ and it will be seen that to expand it we write down first the diagonal from left to right and subtract the diagonal from right to left.

It is clearly the same if we write $\begin{vmatrix} A_1 A_2 \\ B_1 B_2 \end{vmatrix}$ for evaluating this in the same way gives the same expression, but if we interchange either the rows or the columns we obtain the previous expression with a reversed sign. Thus

$$\begin{vmatrix} B_1 A_1 \\ B_2 A_2 \end{vmatrix} = B_1A_2 - A_1B_2 = - \begin{vmatrix} A_1 B_1 \\ A_2 B_2 \end{vmatrix}$$

Determinant notation enables the solution of equations to be written down at once, whereas in the expanded form the expressions would be too cumbersome to remember.

Suppose we have two equations

$$A_1x + B_1y + C_1 = 0$$
$$A_2x + B_2y + C_2 = 0$$

We can solve this at once by writing*

$$\frac{x}{\begin{vmatrix} B_1 C_1 \\ B_2 C_2 \end{vmatrix}} = \frac{y}{\begin{vmatrix} C_1 A_1 \\ C_2 A_2 \end{vmatrix}} = \frac{1}{\begin{vmatrix} A_1 B_1 \\ A_2 B_2 \end{vmatrix}}$$

Note the symmetry of the middle term, i.e. the sequence $ABCA$ is maintained. Otherwise the sign will be wrong as just explained. The simplicity of this form will be apparent particularly if $A$, $B$ and $C$ are not simple numbers but complex expressions.

* The reader may verify this by expanding the determinants and then solving the equations in the usual way, when identical results will be obtained.

As an example let us solve the expressions for the circuit of Fig. (i). The equations, suitably rearranged, are

$$i_1(R_1 + R_3) - i_2R_3 - e = 0$$
$$- i_1R_3 + i_2(R_2 + R_3) = 0$$

Then

$$\frac{i_1}{\begin{vmatrix} -R_3 & -e \\ (R_2 + R_3) & 0 \end{vmatrix}} = \frac{i_2}{\begin{vmatrix} -e & (R_1 + R_3) \\ 0 & -R_3 \end{vmatrix}}$$
$$= \frac{1}{\begin{vmatrix} (R_1 + R_3) & -R_3 \\ -R_3 & (R_2 + R_3) \end{vmatrix}}$$

Whence

$$i_1 = \frac{e(R_2 + R_3)}{(R_1 + R_3)(R_2 + R_3) - R_3^2} = \frac{e(R_2 + R_3)}{R_1R_2 + R_2R_3 + R_3R_1}$$
$$i_2 = \frac{eR_3}{R_1R_2 + R_2R_3 + R_3R_1}$$

The current in $R_3 = i_1 - i_2$

$$= \frac{eR_2}{R_1R_2 + R_2R_3 + R_3R_1}$$

which is the same as was obtained by Thevenin's Theorem.

For a three variable expression we obtain a third order determinant of the form

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix}$$

This can be simplified into

$$A_1 \begin{vmatrix} B_2C_2 \\ B_3C_3 \end{vmatrix} + A_2 \begin{vmatrix} B_3C_3 \\ B_1C_1 \end{vmatrix} + A_3 \begin{vmatrix} B_1C_1 \\ B_2C_2 \end{vmatrix}$$

Note again the sequence of the second term to preserve correct sign. As an example, let us take the equations

$$A_1x + B_1y + C_1z = 0$$
$$A_2x + B_2y + C_2z = 0$$
$$A_3x + B_3y + C_3z = e.$$

We rewrite these, mentally, in a form involving a constant term $D$ such that the equation $= 0$ in each case. In the

first two equations $D = 0$ while in the third it is $-e$. Then

$$\frac{x}{\begin{vmatrix} B_1C_1 & 0 \\ B_2C_2 & 0 \\ B_3C_3 & -e \end{vmatrix}} = \frac{y}{\begin{vmatrix} C_1 & 0 & A_1 \\ C_2 & 0 & A_2 \\ C_3 & -e & A_3 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 0 & A_1B_1 \\ 0 & A_2B_2 \\ -e & A_3B_3 \end{vmatrix}} = \frac{1}{\begin{vmatrix} A_1B_1C_1 \\ A_2B_2C_2 \\ A_3B_3C_3 \end{vmatrix}}$$

which determines $x$, $y$ and $z$. Usually the individual solutions will simplify either directly or on expansion. For example the third determinant resolves itself into

$-e \begin{vmatrix} A_1B_1 \\ A_2B_2 \end{vmatrix}$, the remaining terms going out.

Similarly the first determinant can be expanded directly into

$$B_1 \begin{vmatrix} C_2 & 0 \\ C_3 & -e \end{vmatrix} + B_2 \begin{vmatrix} C_3 & -e \\ C_1 & 0 \end{vmatrix} + B_3 \begin{vmatrix} C_1 & 0 \\ C_2 & 0 \end{vmatrix} = -B_1C_2e + B_2C_1e$$

or we can interchange columns 1 and 3, changing the sign, which gives

$$e \begin{vmatrix} C_1B_1 \\ C_2B_2 \end{vmatrix} = C_1B_2e - B_1C_2e, \text{ as before.}$$

With third- and higher-order determinants there are often other simplifications, based on certain simple rules. The first we have already seen, viz.

**1. Interchanging two columns (or rows) of a determinant changes its sign.**

Other rules are—

**2. Multiplying any column (or row) by a constant multiplies the whole determinant by that factor.**

Thus $\begin{vmatrix} mA_1B_1 \\ mA_2B_2 \end{vmatrix} = m \begin{vmatrix} A_1B_1 \\ A_2B_2 \end{vmatrix}$

as the reader may easily verify.

**3. If each constituent of any column or row comprises two terms the determinant may be expressed as the sum of two simple determinants.**

Thus $\begin{vmatrix} (A_1 + a_1) & B_1 \\ (A_2 + a_2) & B_2 \end{vmatrix} = \begin{vmatrix} A_1B_1 \\ A_2B_2 \end{vmatrix} + \begin{vmatrix} a_1B_1 \\ a_2B_2 \end{vmatrix}$

which again may be simply verified.

**4. If the constituents of any column or row are increased or diminished by equal multiples of another column or row the result is unaltered.**

Thus $$\begin{vmatrix} (A_1 + mB_1) & B_1 \\ (A_2 + mB_2) & B_2 \end{vmatrix} = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}$$

which is often useful in effecting simplifications.

# ANSWERS TO EXAMPLES

I. (1) (a) $2\mu H$. (b) 80%.
   (3) $L_1 = 40\mu H$. Tap approximately 0·45 up the coil.

III. (1) 547 ohms. (2) $\frac{1}{8}$th tap.

IV. (1) (i) $64{\cdot}5\,\mu H$; 3 : 1. (ii) $80\mu H$.; 2·5 : 1.
   (2) 5 ohms at 600 kc.; 27·7 ohms at 1 200 kc.
   (3) 60 at 600 kc.; 92 at 1 200 kc.

V. (1) (a) $183\mu H$.; 222 000 ohms; (c) 444; (d) one-half.
   (2) 111.

VI. (1) 5·6 volts; 1·2 volts (r.m.s.) (2) 2·2 volts.

VIII. (1) $21\mu H$.; 5·55. (2) 24 kc.; 18 kc.; 12 kc. No difference.
   (3) $0{\cdot}0104\mu F$.; 12 kc. at 1 200 kc., 24 kc. at 600 kc.

X. (1) $r = 490{\cdot}9$ ohms, $s = 121{\cdot}2$ ohms.
   (2) Series arm. Resistance of 180 000 ohms shunted by condenser of $0{\cdot}00363\mu F$.
   Shunt arm. Resistance of 2 220 ohms in series with 1·45H.
   (3) See Fig. 101.

# EXAMINATION PAPERS

By courtesy of the City and Guilds of London Institute (Department of Technology), we are able to reproduce the papers set in the 1945 and 1946 examinations. Some of the questions deal with matters which are discussed in *Short-wave Radio* or *Cathode Ray Oscillographs*.

In each case only six of the ten questions were required to be answered.

## 50c. RADIO-COMMUNICATION (1945)

### Grade II

1. Describe the principles underlying the operation of an oscillator in which the frequency is determined by a resistance-capacitance network.

2. What is the purpose of—

(*a*) neutralizing,

(*b*) decoupling,

as applied to thermionic valve circuits?

How are these operations carried out?

3. Describe the smoothing circuit of a power pack of a receiver working from a.c. electric supply mains. Show how to estimate values of inductance and capacitance involved in the anode supply smoothing circuit.

4. A receiver tuned to a remote transmitting station on a frequency of 175 kc/s suffers interference from a nearby transmitting station working on 160 kc/s.

By a suitable combination of inductance and capacitance in the aerial circuit of the receiver, the interference is reduced. Describe such an arrangement and calculate the approximate values that should be used.

5. Explain the method of operation of the cathode ray tuning indicator (magic eye) and state in what part of the circuit of a receiver it should be connected.

6. Two circuits similar in all respects and consisting of a condenser and resistance are so arranged that when one is being charged the other is being discharged, the same source of voltage being used for charging each condenser in turn. The commencement of charging one condenser and discharging of the other condenser occurs at the same moment. If the time constant of each circuit is one second, what time elapses before the voltages across the condensers in both circuits are equal?

7. Give a diagram of a master oscillator circuit and explain how constancy of the frequency generated is obtained.

8. Explain the meaning of "depth of modulation."

The depth of modulation of a carrier is reduced from 60 per cent to 30 per cent; what is the ratio of the total power radiated in the two cases?

9. How could an appreciable reduction be effected in interference with radio reception due to—

(*a*) mains hum,

(*b*) medical appliances, e.g., therapy apparatus?

10. Describe a method of measurement of a radio frequency field strength of the order of several micro-volts per metre.

## 50c. RADIO-COMMUNICATION (1945)

### Grade III

1. In one form of Adcock direction finder, the movable coil of the radiogoniometer is rotated uniformly by an electric motor and the output connected to a cathode ray tube. When the receiver is tuned to a distant station a trace appears on the tube.

Give a diagram of connections to the cathode ray tube, explain how the time base can be obtained, and draw a sketch showing the approximate shape of the trace.

2. Give an outline of the principles of propagation along a wave guide transmission line at radio frequencies.

3. The audio-frequency output of a superheterodyne receiver is fed through a cable pair to a remote control station. Using the same pair as a control line, the remote operator is required to vary the frequency of the receiver oscillator slightly to keep the set in tune. Describe a suitable system, giving particulars of the connections at both ends of the pair and the devices used at both ends of the circuit.

4. Eight similar vertical wires of a transmitting array fed in phase are placed in a row half a wave length apart at the same height above ground. Deduce a formula for the polar diagram in the horizontal plane. What would be the maximum gain of the array as compared with a single similar vertical wire?

5. Give a brief explanation of a frequency discriminator and explain how it can be used to control the frequency of a valve oscillator.

6. "The earth's magnetic field renders the ionosphere doubly refracting to radio waves." Discuss this statement.

7. What is meant by "the dynamic resistance of a tuned circuit"?

A parallel circuit consists of two arms, one having inductance and resistance in series, and the other having capacitance. Deduce an expression for the dynamic resistance of the circuit at resonance.

8. Explain, with outline diagrams, how suppressed carrier single sideband telephony radio emissions are obtained. What are the advantages of this type of emission?

9. If $f_r$ is the resonant frequency of a tuned circuit and $f_2$ and $f_1$ are the frequencies above and below $f_r$ at which the current in the circuit is $\frac{1}{\sqrt{2}}$ of the current at resonance, the applied voltage being constant, prove the formula for *either* the decrement *or* the Q-factor.

10. A two-wire land line is connected to a four-wire circuit for radio telephony transmission and reception. Give a block diagram and a brief explanation of the circuit arrangement and devices required to effect voice switching of the transmit-receive circuit.

# 50c. RADIO-COMMUNICATION (1946)

## Grade II

1. Upon what factors do the selectivity and sensitivity of a superheterodyne receiver depend?

2. Describe briefly two methods of speech amplitude modulation in a radio-frequency transmitter.

3. A vertical aerial without a top has the same length as the vertical members of a rectangular single turn loop aerial. It is found that the signal from a transmitting station is stronger on the vertical aerial than on the loop when the latter is in a position for maximum reception. Explain the reasons.

4. Describe briefly two methods of determining modulation factor by means of a cathode ray oscillograph.

5. The inductances of two inductively coupled coils, $A$ and $B$, are 0·002 and 0·032 henrys respectively, and the coupling coefficient is 0·75. Regarding the resistance of the coil $A$ as negligible, what will be the open circuit voltage across the terminals of coil $B$ when 0·5 volt is applied across the terminals of $A$?

6. Describe the action of a crystal resonator, giving the equivalent circuit and a curve showing the approximate relationship between impedance and frequency.

7. Explain how control of focus is achieved in a cathode ray tube by electrostatic means.

8. A single stage thermionic valve amplifier has a tuned anode circuit having a capacitance of 400 micro-micro-farads, an inductance of 300 micro-henrys, and a resistance of 12 ohms. If the plate resistance is 40,000 ohms and the amplification factor is 10, what is the radio frequency voltage across the tuned output circuit when 0·12 volt is applied to the grid of the valve?

9. Describe the causes of the various classes of noise encountered in a radio receiver.

10. What is the purpose of *delayed* automatic gain control in a receiver? Explain how it is obtained.

## 50c. RADIO-COMMUNICATION (1946)

### Grade III

1. Distinguish clearly between amplitude, phase, and frequency modulation.

2. In the carrier shift radio telegraph system two carrier frequencies, separated by about 800 cycles, are used—one for mark and one for space. Describe some existing device, or outline a possible device of your own, for keying the transmitter to ensure that the difference between mark and space carrier frequencies remains substantially constant.

3. In a three-channel amplitude modulated radio telegraphic system, six tones are used, three for mark and three for space. Assuming the maximum speed of transmission to be 25 reversals per second, what tone frequencies would you use? Give reasons for your answer.

Describe the method of separating the received tones.

4. Describe, with constructional particulars, the ring type valve seal, and discuss its advantages.

5. (i) Explain the meaning of Class A, B, and C amplification.

(ii) What order of efficiency could be expected from a large power amplifier, employing the three different classes? Give reasons for the higher efficiency of one class as compared with another.

6. What is the meaning of d.c. modulation as applied to a radio television transmission? Explain the advantages of this type of modulation.

7. Explain the necessity of correctly matching a transmission line to an aerial. A high frequency transmitter is connected to the middle of a horizontal dipole, consisting of two members each a quarter of a wave-length long, by means of an open wire feeder. Describe, with the necessary calculations, a type of matching unit suitable for connecting the feeder to the dipole.

8. Describe a method of measuring the harmonic distortion of an audio-frequency amplifier.

9. Discuss the effect of the dielectric constant and conductivity of earth upon the radiation pattern (polar diagram) in a vertical plane, due to—

(*a*) a vertical short wave dipole;
(*b*) a horizontal short wave dipole.

10. A medium frequency signal generator has a variable attenuator in its output circuit. Give an outline of a design of the attenuator to feed into an unbalanced output of 70 ohms.

# INDEX

| | | |
|---|---|---|
| Duration of Loan | Students/Spl/'C'<br>Text Books — 3 days<br>Technical Books — 7 days<br>General Books — 14 days | Teachers—'A'<br>One month |

FROM THE DATE OF ISSUE